ACCENT IN CARRANZA TSOTSIL[1]

Gilles Polian and Eduardo de Jesús López Mendoza

March, 2024
To appear, International Journal of American Linguistics

1. Introduction

Few Mayan languages have developed phonological tone; confirmed cases include Yucatec Maya, Uspanteko and Mocho' (see Bennett 2016, Campbell 2017, England and Baird 2017 and references therein). Additionally, a controversy exists about the presence of tones in Carranza Tsotsil (henceforth CT; a.k.a San Bartolo Tsotsil).[2] On the one hand, Sarles (1966) described two tones in CT, H and L (high and low), as in the minimal pair /ʔólìl/ 'child' vs /ʔòlíl/ 'half' (*ibid.*, p.28), displaying contrastive HL vs LH pitch patterns. Sarles also observed that H and L tones typically alternate on every other syllable, as in /hpímùbtéswànéh/ HLHLH '(the) person who makes it thick' (*ibid.*, p.30). Kaufman (1972:84-88) endorsed Sarles' proposal and explained the emergence of an initial L tone out of the loss of */h/ in V_C contexts (V=vowel, C=consonant). Concretely, *CVhC sequences became CV̀C, for instance **pohp* > *pòp* 'mat'. On the other hand, Herrera (2013, 2014) disputes the existence of phonological tones in CT. For instance, she shows that in her data *pòp* 'mat' displays no low or falling f0 (fundamental frequency). Avelino *et al.* (2011: 1) also report that "fieldwork by Avelino on San Bartolo Tzotzil in 2010 did not reveal reliable tone contrasts".

In this paper, we claim that Sarles (1966) and Kaufman (1972) were right about CT prosody, at least at the phonetic level, and we explain the divergence between our data and Herrera's data as the impact of phrasal intonation on word-level prosody. Based on a quantitative and qualitative examination of firsthand data, we confirm the existence of a contrast between prosodically prominent vs non-prominent syllables in CT, a contrast that we first characterize in terms of *accent*, where accented/unaccented matches Sarles' H/L tone respectively. Refining Sarles' preliminary description, we argue that word-level prosodic patterns underly an opposition between two lexical accent classes and describe several phonological phenomena that affect accent. On the analytical level, we propose that CT accent results from the interaction between tone and stress: some morphemes bear a L or H phonological tone, which affects the basic trochaic metrical pattern of the language through the known affinity between H tones and foot heads on the one hand and between L tones and foot non-heads on the other hand (de Lacy 2002). In the end, we recast the

[1] We would like to thank the inhabitants from Paraíso del Grijalva, Venustiano Carranza, for sharing with us their linguistic knowledge, especially Juan Carlos López de la Torre, Daniela López Mendoza and Miguel Angel Solano Hidalgo. We are also grateful to our fellow linguists Esther Herrera, Roberto Zavala, Eladio Mateo and Néstor Hernández for their comments and Mario Chávez for reading an early draft of this paper and always being willing to discuss phonology issues with us and give us useful advice. We also thank the two IJAL reviewers (Ryan Bennett and an anonymous reviewer), David Beck and an IJAL associate editor, who provided very helpful comments and criticisms. We are especially indebted to Ryan Bennett for providing decisive suggestions for improving the methodology and analysis and for bringing to our attention literature we were unaware of. All shortcomings remain our own responsibility.

[2] Previously spelled *Tzotzil*, ISO 639-3 tzo.

opposition between accented vs non-accented syllables in tonal terms as H-toned vs non-H-toned syllables.

2. Background

Carranza Tsotsil, also known as San Bartolo Tsotsil, is a western Mayan language spoken by approximately 10 000 people in Venustiano Carranza, previously named San Bartolomé de los Llanos, a municipality in the lowlands of Central Chiapas, Mexico. Our data come specifically from Paraíso del Grijalva, a village located in the southern part of the municipality. No internal dialect variation is known to Carranza. CT is the southernmost dialect of Tsotsil, a language of which only one dialect, Zinacantec, has been well-described, at least concerning its lexicon (Laughlin 1975) and morphosyntax (Haviland 1981; Aissen 1987; *inter alia*). Tsotsil is closely related to Tseltal (ISO 639-3 tzh) and together they form the Tseltalan branch (Polian 2017).

Few linguistic studies exist on CT. Sarles (1966) is a preliminary phonological and grammatical description. Kaufman (1972) devotes a few pages to its phonological and lexical dialectal features. Finally, Herrera (2013, 2014) explores its phonology: she studies the phonetic properties of vowels, main consonantal allophones, syllable structure and accentual patterns.

The CT phonemic inventory is provided in Tables 1 and 2, with the practical orthography we adopt here indicated in angle brackets when it differs from the International Phonetic Alphabet. Like all Mayan languages, CT has series of plain and ejectives stops and affricates. It also has an implosive /ɓ/, realized as a laryngealized nasal in coda position, which we represent as <m̰> (e.g. /ʧaɓ/ <cham̰> ‘wax’).[3]

TABLE 1

CONSONANTS

	Labial	Alveolar	Palato-alveolar	Palatal	Velar	Glottal
Stop	p p’ ɓ <b, m̰>	t t’			k k’	ʔ
Affricate		ts ts’	ʧ <ch> ʧ’ <ch’>			
Fricative	β <v>	s	ʃ <x>		x <j>	
Sonorant	m	n l ɾ <r>		j <y>		

TABLE 2

VOWELS

	Front	Central	Back
High	i	ɨ <u>	
Mid	e		o
Low		a	

CT is strictly quantity-insensitive: it has no vowel length, no diphthongs, nor any phonological phenomenon sensitive to syllabic weight. Syllables are predominantly CV(C) and most lexical roots follow a CVC or CV.CVC syllabic pattern, with CV, CV.CV, CV.VC and CVC.CVC as

[3] We follow official orthographic recommendations from the Instituto Nacional de Lenguas Indígenas (INALI) for Tsotsil, except for the glottal stop and the <m̰> allophone of /ɓ/. In the rest of the paper, we use exclusively the practical orthography to transcribe CT data.

secondary patterns, especially among nouns. There are no restrictions regarding which consonant occupies the onset or coda position in the syllable.

Concerning word stress, Herrera (2014) describes a unique prominence on the final syllable. We show in §3.4 that this final prominence pertains to the phrasal level. Instead, we suggest that in CT default word stress is, in fact, initial (§6.3).

3. Accented and unaccented syllables

CT word-level prosody follows specific patterns, in which some syllables are pronounced with a higher pitch than others, and often also with a greater intensity and/or a longer duration. Adopting "accent" as a preliminary descriptive term for this phenomenon, we will call "accented" and mark with an acute accent the syllables that display this kind of prominence, while the comparatively less prominent syllables will be designated as "unaccented" and written with a grave accent. Although in our analysis "accented" will ultimately be replaced by "H-toned", we cannot take tone for granted from the outset, since what is at stake is the mere existence of tone. We present in Table 3 the most important generalizations we've formulated. Several key phenomena stand out: two accent classes (words that follow the default prosodic pattern vs those that display a marked pattern), deaccenting, resulting from a ban on adjacent accented syllables and the special prosodic status of the last syllable of intonational phrases.

TABLE 3

SIX GENERALIZATIONS ON CT ACCENTUATION

	Patterns	Examples
I. Default alternating pattern: odd-numbered syllables are accented, even-numbered syllables are unaccented. Words that follow this pattern = "class 1".	σ́ σ́.σ̀ σ́.σ̀.σ́ σ́.σ̀.σ́.σ̀ σ́.σ̀.σ́.σ̀.σ́ etc.	[chóy] 'fish' [ví.nìk] 'man' [ʔí.nì.tám] 'iguana' [jpóx.tà.vá.nèj] 'doctor' [jchá.nùm̥.tés.và.néj] 'teacher'
II. Marked pattern: the 1st syllable is unaccented. - Disyllables: the 2nd syllable is accented. - Longer words: both the 1st and 2nd syllables are unaccented; the default alternation is resumed from the 3rd syllable onwards. Words that follow this pattern = "class 2".	σ̀ σ̀.σ́ σ̀.σ̀.σ́ σ̀.σ̀.σ́.σ̀ σ̀.σ̀.σ́.σ̀.σ́ etc.	[pòp] 'mat' [bò.lóm] 'jaguar' [mà.tà.nál] 'gift' [tòj.bà.lá.lìl] 'employee' [tòj.bà.lá.lì.lún] 'I am an employee'
III. Class-shifting suffixes: some suffixes (-σ$^{>II}$) force words into marked pattern II.	 **σ́**+-σ$^{>II}$→ **σ̀**.σ́ **σ́**.σ̀+-σ$^{>II}$→ **σ̀**.σ̀.σ́ **σ́**.σ̀.σ́+-σ$^{>II}$→ **σ̀**.σ̀.σ́.σ̀	Passive suffix *-ot*$^{>II}$: pás → [pà.sót] 'be done' élk'àn → [èl.k'à.nót] 'be stolen' chíkìltá → [chì.kìl.tá.òt] 'be tickled'
IV. Deaccenting: in a sequence of contiguous accented syllables, all but the last one get deaccented.	σ́+σ́ → σ̀.σ́ σ́+σ́+σ́ → σ̀.σ̀.σ́ etc.	ʔóy+chóy→[ʔòy.chóy] 'there is fish' ʔóy+jún+chóy→[ʔòy.jùn.chóy] 'there is one fish'
V. Accented suffixes: some long suffixes (2 or 3 syllables) have a pre-defined accent on a specific vowel (taking precedence over	σ́+-σ́.σ̀ → **σ̀**.σ́.σ̀ σ̀+-σ́.σ̀ → **σ̀**.σ́.σ̀ σ̀.σ́+-σ̀.σ́ → **σ̀**.**σ̀**.σ̀.σ́	jchóy+-tútik→[jchòy.tú.tìk] 'our fish' jpòp+-tútk→[jpòp.tú.tìk] 'our mat' bòlóm+-etík→[bò.lò.mè.tík] 'jaguars'

patterns I and II, provoking deaccenting following IV).		
VI. Final intonation: the last syllable of intonational phrases is neutralized for accent (symbolized here as “σ$^{\times}$”); its prosody conveys sentence level pragmatic meanings; no pre-deaccenting (IV) occurs before that syllable.	σ́.σ̀]$_{\text{IP}}$→ σ́.σ$^{\times}$ σ̀.**σ̀**.σ́]$_{\text{IP}}$→ σ̀.**σ́**.σ$^{\times}$ σ́+σ́]$_{\text{IP}}$→ σ́.σ$^{\times}$	ví.nìk]$_{\text{IP}}$ → [ví.nik$^{\times}$] ‘man’ mà.tà.nál]$_{\text{IP}}$ → [mà.tá.nal$^{\times}$] ‘gift’ ʔóy+chóy]$_{\text{IP}}$ →[ʔóy.choy$^{\times}$] ‘there is fish’

3.1. Accent classes 1 and 2

Prosodic words in CT show a binary opposition between those that start with an accented syllable and follow the default alternating pattern of I and those that start with an unaccented syllable and fit into the marked pattern of II. We respectively label these two word types as “(accent) class 1” and “(accent) class 2” words. An example of the default accent alternation is given with the inflected forms of *pój* ‘take away’ (class 1) in (1a) and (1b): as an additional syllable is added with the plural suffix *-tan* in (1b), one passes from a four-syllable pattern [σ́.σ̀.σ́.σ̀] to one of five syllables [σ́.σ̀.σ́.σ̀.σ́]. Observe how the suffixes *-oj* ‘perfect’ and *-talel* ‘directional’ get opposite accents as compared to (1a).[4]

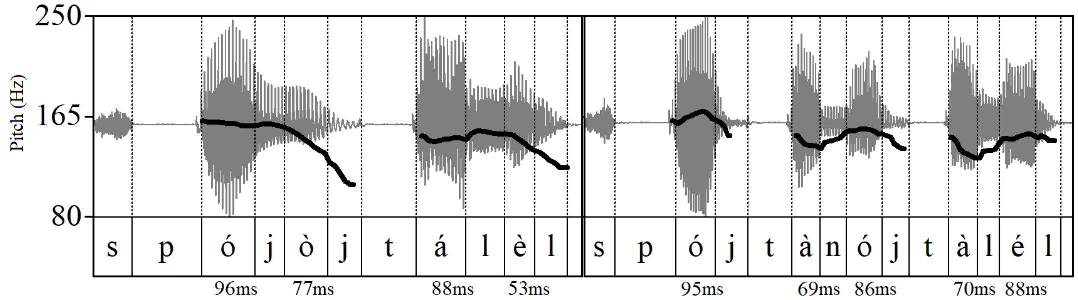
(1)a.	Lìʔ	**s–pój–òj–tálèl**	tà	tèklúm=e.[5]
	here	3ERG–take.away–PERF–DIR	PREP	Carranza=DEF
	‘She has taken it away (from someone) and brought it here to Carranza.’			
b.	Lìʔ	**s–pój–tàn–ój–tàlél**	tà	tèklúm=e.
	here	3ERG–take.away–PL–PERF–DIR	PREP	Carranza=DEF
	‘She has taken them away (from someone) and brought them here to Carranza.’			

Figure 1 gives and idea of the phonetic manifestation of this accentual contrast in terms of pitch and duration: accented syllables are associated with a high or rising pitch and a longer duration, unaccented syllables with a low or falling pitch and a shorter duration. We’ll see below that intensity can also be a cue for accent and confirm these claims with quantitative data. In this and all subsequent figures, the y-axis scale in Hz on the left corresponds to the pitch track, while the waveform is normalized from 1 to -1. The x-axis measures time, but only relevant vowels are measured (in milliseconds).

[4] Abbreviations: 1, 2, 3: 1st, 2nd, 3rd person; ABS: absolutive; DEF: definite; DEM: demonstrative; DIR: directional; ERG: ergative; EXC: exclusive; INC: inclusive; IPFV: imperfective; IRR: irrealis; NEG: negation; PERF: perfect; PFV: perfective; PL: plural; POSS: possessive; PREP: preposition.

[5] As highlighted in Table 3, row VI, the last syllable of an intonation phrase is neutralized for accent. In all examples, we leave that syllable without a graphic accent.

Figure 1: Waveform, pitch track and vowel duration from the production of *spójòjtálèl* ‘she has taken it away’ (left, from (1a)) and *spójtànójtàlél* ‘she has taken them away’ (right, from (1b)) by speaker M1, showing the alternating pattern of accented and unaccented syllables



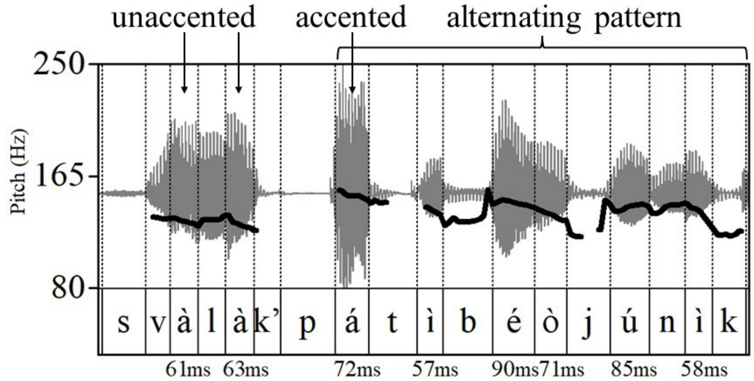
The most notable characteristic of class 2 words is their unaccented first syllable, even if it is the only syllable of a monosyllabic word. Additionally, the second syllable is also unaccented in words of three or more syllables. Note that from the third syllable onward, class 1 and 2 are identical, with an accent on every odd-numbered syllable. The fact that all class 2 words follow pattern II as defined in Table 3 can be verified through suffixation, as most suffixes add a syllable to their stem without altering the established patterns. This is done in (2) by adding suffixes to the class 2 intransitive root *nòp* ‘get used to’, producing stems of increasing length. Note for instance that the suffix *-ik* ‘plural’ is accented in (2b) because it makes up the second syllable in the pattern [σ̀.**σ́**], whereas in (2d) it is unaccented, as expected from its position as the fourth syllable of the pattern [σ̀.σ̀.σ́.**σ̀**].

(2)a.	nòp ‘get used to’		→ [nòp]
b.		+ -ik ‘plural’	→ [nòpík]
c.		+ -ilaj ‘iterative’	→ [nòpìláj]
d.		+ -ilaj ‘iterative’ + -ik ‘plural’	→ [nòpìlájìk]

An 8-syllable class 2 word is shown in (3) and Figure 2. Not all syllables are readily comparable here because of the varying vowel qualities and syllable types, but the contrast between the unaccented first two syllables and the accented third syllable can at least be appreciated in terms of pitch and intensity (visible through the amplitude of the waveform). Also of note is the alternating accent from the third syllable onward.

(3) S-vàlàk’pátì-bé-òj-ún-ìk lìʔ=e.
3ERG-turn.one’s.back.on-APPL-PERF-1ABS-PL=DEF here=DEF
‘They have affected me by turning their backs on her here.’

Figure 2: Waveform, pitch track and vowel duration from the production of *svàlàk’pátìbéòjúnìk* ‘they have affected me by turning their backs on her’ from (3) by speaker M1, illustrating an 8-syllable class 2 word



Class 2 can be considered as marked with respect to class 1 in several respects: historically, morphologically and in terms of frequency. Kaufman (1972) established its main historical source: most class 2 words come from proto-Tseltalan forms with a laryngeal fricative */h/ in V_C contexts, namely *CV**h**C and *CV**h**CVC.[6] According to Kaufman, the loss of */h/ in this context was compensated by the emergence of a L tone. What is clear is that whatever makes a class 2 word deviate from the default pattern is largely a continuation of */h/ and, as such, it keeps carrying out the same functional load in the lexicon and grammar. In Tseltal (the sister language of Tsotsil), the /h/ in V_C contexts can be lexical (as an unanalyzable part of the root) or derivative. For instance, it derives an intransitive anticausative (inchoative) from transitive CVC roots (e.g. *jam* ‘open’ → *jahm* ‘become open’), and appears in many other derivations combined with different suffixes (e.g. <*h*> + -*Vb* derives an instrument noun: *jahmib* ‘opener’). The same can be said of pattern II in CT: it is lexical in some cases and derivative in others, alone or in combination with suffixes: *jám* (class 1) ‘open’ → *jàm* (class 2) ‘become open’ → *jàmòbíl* (class 2) ‘opener’. Accordingly, the direction of the derivation is always from class 1 to 2, as some suffixes (class-shifting suffixes, cf. generalization III of Table 3) force words into class 2, regardless of the base’s accent class (i.e., they also derive class 2 stems from class 2 roots), never the other way around. Finally, just as roots that start with CVhC in Tseltal are less frequent than those that start in CVC, in CT class 1 roots outnumber class 2 roots. The distribution by accent class varies according to the lexical category of the root: all transitive and positional roots (several hundreds of items) are class 1, as they stem from Proto-Mayan *CVC roots, while other categories are mixed. In our current database, the percentage of class 2 items with nominal/adjetival roots is 29% (from a total of 382 items) and goes up to 47% for the few intransitive roots (total: 38).

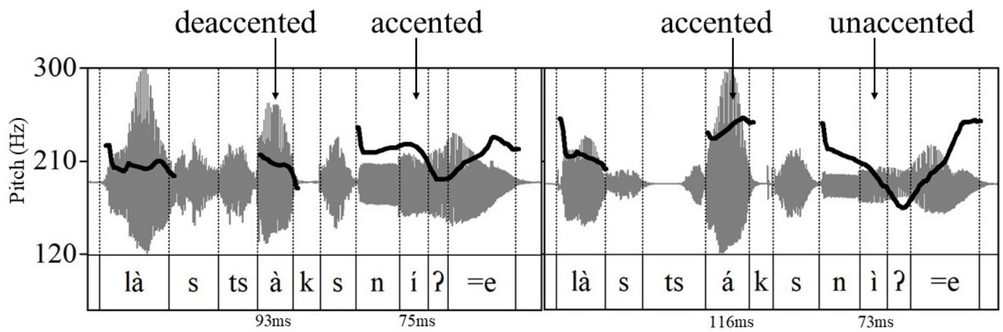
3.2. Deaccenting

The anticipatory deaccenting in sequences of potentially accented syllables is best illustrated by a minimal pair such as in (4): before the accented *sníʔ* ‘her nose’, the transitive verb *tsák* ‘grab’ becomes unaccented, (4a), whereas it remains accented before the unaccented (class 2) *snìʔ* ‘her son-in-law’, (4b); see Figure 3 for a pronunciation sample.

[6] Preconsonantal /h/ never appeared outside of root-initial syllables in proto-Tseltalan, e.g. no *CVCV**h**C form existed.

(4)a. Là s–ts**à**k s–n**í**ʔ=e.
PFV 3ERG–grab 3POSS–nose=DEF
‘She grabbed her nose.’
b. Là s–ts**á**k s–n**ì**ʔ=e.
PFV 3ERG–grab 3POSS–son-in-law=DEF
‘She grabbed her son-in-law.’

Figure 3: f0 and vowel duration for ‘she grabbed her nose’ (left) and ‘she grabbed her son-in-law’ (right) by speaker F



The same effect obtains with polysyllables: class 1 words cause pre-deaccenting, (5a), while class 2 words don’t, (5b), with the minimal pair *ʔólìl* ‘child’ - *ʔòlíl* ‘half’. Note that this differential prosodic effect on the preceding syllable comes as a useful ancillary clue for discriminating both accent classes.

(5)a. Là s–ts**à**k **ʔó**lìl=e.
PFV 3ERG–grab 3POSS–child=DEF
‘She grabbed the child.’

b. Là s–ts**á**k **ʔò**líl=e.
PFV 3ERG–grab half=DEF
‘She grabbed half.’

Deaccenting affects whole strings of potentially accented syllables: all but the last one get deaccented. This allows us to distinguish a deaccented syllable from an unaccented one: although they both lack an accent, the former but not the latter keeps its pre-deaccenting effect. This can be verified by adding an accented syllable such as *té* ‘there’ after the minimal pair of (4): observe in (6a) how *sníʔ* ‘her nose’ is deaccented to *snìʔ*, and thus appears as locally neutralized with *snìʔ* ‘her son-in-law’, while still deaccenting the previous verb *tsák* ‘grab’. In contrast, in (6b) the verb is not deaccented before the inherently unaccented ‘son-in-law’.

(6)a. Là s–ts**à**k s–n**ì**ʔ **té**=e.
PFV 3ERG–grab 3POSS–nose there=DEF
‘She grabbed her nose there.’

b. Là s–ts**á**k s–n**ì**ʔ **t**é=e.
PFV 3ERG–grab 3POSS–son-in-law there=DEF
‘She grabbed her son-in-law there.’

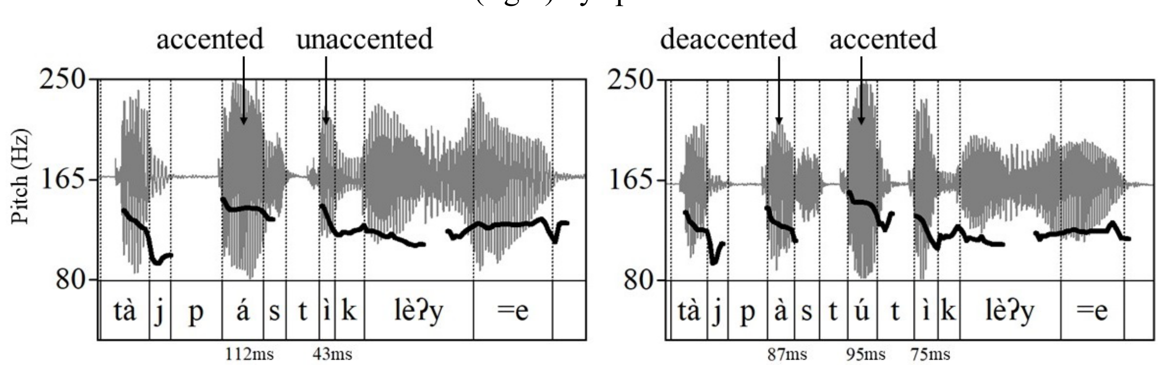
This phenomenon confirms the lexical nature of the prosodic difference between Class 1 and 2: nothing in the supra-lexical intonation or in the information structure could explain the fact that the verb is deaccented in (6a) but not in (6b).

3.3. Accented suffixes

Most suffixes add at least one syllable to the stem; long suffixes add two or three. None of the monosyllabic suffixes modifies the default alternation between accented and unaccented syllables, as they just take the opposite accent value of the preceding syllable (see (1) above).[7] In contrast, some long suffixes, especially (but not only) plural morphemes, have a pre-defined accent on a specific vowel, such as *-tútik* ‘plural of 1st person exclusive’, *-etík* ‘nominal plural’, *-ukútik* ‘1st person plural inclusive absolutive’,[8] *-ebál* ‘be about to X’, among others. When this accent falls in a position that is normally unaccented, that is, in an even-numbered syllable, it takes precedence over the default accent alternation. When it appears just after an accented syllable, the latter is deaccented, as expected by the effect seen in §3.2. Compare (7a), where the inclusive suffix *-tik* on the class 1 root *pás* ‘do’ forms a regular disyllabic [σ́.σ̀] word, with (7b): here, the accented exclusive suffix *-tútik* creates a [σ̀.σ́.σ̀] word by forcing an accent on the second syllable and deaccenting the root. Figure 4 illustrates the phonetics of this contrast.

(7)a. Tà j–**pás–tìk** lèʔy=e.
IPFV 1ERG–make–PL.INC DEM=DEF
‘We (inclusive) make this one.’
b. Tà j–**pàs–tútìk** lèʔy=e.
IPFV 1ERG–make–PL.EXC DEM=DEF
‘We (exclusive) make this one.’

Figure 4: f0 and vowel duration for ‘we (incl.) make this one’ (left) and ‘we (excl.) make this one’ (right) by speaker M1



[7] Sarles (1966:29) describes some monosyllabic suffixes as inherently H-toned (passive *-ot*, causative *-tes*), but we claim that this is inaccurate.
[8] Even though these three plural suffixes are historically segmentable in smaller morphemes (an element *tik* is for instance shared here), they function synchronically as morphologically unanalyzable units.

No suffix appears to have a pre-defined *unaccented* syllable. In a sense, the situation is the opposite to that of lexical roots: marked roots come with an unaccented (first) syllable, while marked suffixes have an unpredictable accent somewhere. Note also that, unlike class-shifting suffixes (cf. Table 3, generalization III), the accented suffixes don't change the accent class of the stem they attach to: they only alter the default alternance of accents. Anecdotally, some suffixes are at the same time accented and class-shifting, such as *-ebál* 'be about to X': *tál* 'come' (class 1) → ***tà**l-èbál* 'be about to come' (class 2).

3.4. Phrasal intonation

As reported in some other Mayan languages, phrasal prosody in CT is characterized by final prominence, which can eclipse word-level prosody. For instance in Q'anjob'al, Mateo Toledo (2008: 94) states that words in phrase-final position are stressed on the last syllable, while in a non-final position they are stressed on their first syllable. A similar situation has been described for several dialects of Tseltal (Aguacatenango: Kaufman 1971; Petalcingo: Shklovsky 2011; Oxchuc: Polian 2013) and Tsotsil (Zinacantan: Haviland 1981; Huixtan: Cowan 1969).[9] In Haviland's (1981: 14) words, "the primary word stress in Tzotzil falls on the first root syllable; in addition, there is an even stronger stress that occurs on the last syllable of a phrase or sentence (it can consist of a single word pronounced in isolation)".

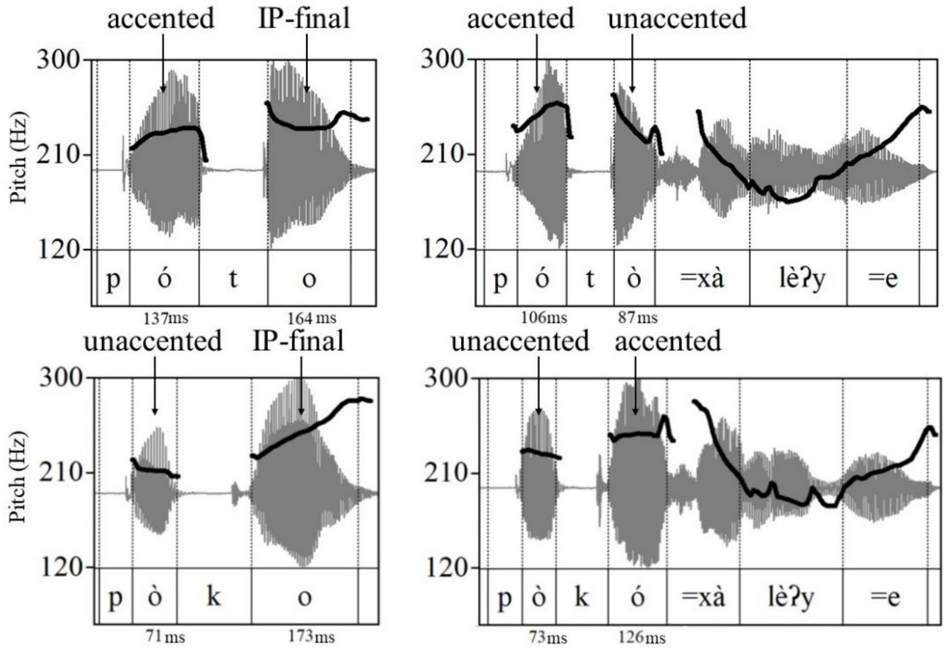
The same applies to CT, where this final prominence at least clearly affects the last syllable of intonational phrases (hereafter IPs), including isolated words, which constitute IPs by themselves (Gordon 2011, 2014; van der Hulst 2014). An illustration of how this modifies accentual contrasts is illustrated in Figure 5 with the near minimal pair *pótò* 'guava' (class 1) vs *pòkó* 'old' (class 2). Their pronunciation in isolation (to the left) displays in both cases a final prominence in terms of greater pitch, intensity and duration. Only in a non-final position (to the right, corresponding to (8)) can one fully appreciate their contrasting accent pattern (σ́.σ̀ and σ̀.σ́ respectively). In fact, despite the final prominence seen in isolation, we'll show in §6.2 below that they maintain their prosodic contrast even in that context, which is already somewhat evident looking at Figure 5, as the asymmetry in prominence between the first and second syllable is much stronger for *pòkó* than it is for *pótò*. On the contrary, we'll argue in §5.5 that monosyllables are completely neutralized for accent in IP-final position.

(8)a. Pótò=xà lèyʔ=e.
guava=already DEM=DEF
'This one is already a guava.'

b. Pòkó=xà lèyʔ=e.
old=already DEM=DEF
'This one is already old.'

Figure 5: Waveform, pitch track and vowel duration from the production of *pótò* 'guava' (upper-left) and *pòkó* 'old' (lower-left) in isolation, showing final prominence at the end of the IP, and in the carrier sentence of (8a) 'this one is already a guava' (upper-right) and (8b) 'this one is already old' (lower-right), without the IP-final impact on the target word, by speaker F

[9] Only Tenejapa Tseltal has been described with final word stress (Berlin 1962, Brown 1996), probably an innovative pattern.



Recall that Herrera (2014) characterized word stress as being final in CT. We argue that this final prominence pertains to IPs, not to words. It means that studying word-level prosody in CT requires the target word to be followed by at least another syllable in the same IP, to act as a buffer which keeps it safe from phrasal intonation. Furthermore, this additional syllable must be unaccented, to avoid pre-deaccenting (cf. §3.2).

3.5. Summary

In summary, CT presents a prosodic contrast between accented and unaccented syllables, with a default binary alternation between both types of syllables in prosodic words. Departures from the default pattern are the result of deaccenting (ban on contiguous accents), phrasal intonation (on the IP-final syllable) or from a lexical or morphological deviation: initial unaccented syllable of class 2 words on the one hand, accented syllable of some long suffixes on the other.

4. Tonal approach and production task

In §5 and §6, we show the results of a production task that corroborate our account of word-level prosody in CT and allow us to assess the plausibility of a tonal approach in line with Sarle's (1966) description, by which the accented/unaccented contrast corresponds to H and L tones respectively. Concretely, if the phonetics of accent contrasts includes a systematic and significant variation in pitch (which is not attributable to phrase-level or other intonational tones), that could count as evidence for phonological tone. Pitch differences should be above the "just noticeable difference" (JND) for lexical tone, which according to Kuang (2013: 3) "appears to be not less than 9 Hz (...), and languages usually require a larger difference than the JND to maintain a phonological contrast; a 20- to 30-Hz difference (2–3 semitones in a comfortable pitch range for males from 150 to 170 Hz, 2.1 st, or for females from 200 to 230 Hz, 2.4 st) is just marginally sufficient [Hart, 1981]". Kuang (2013) cites Siswati (Bantu) and Kiowa (Tanoan), other languages with two level tones, in which pitch differences between H and L tones were reported with a mean difference of 18 and 22Hz respectively (from Maddieson 1978: 339). In terms of semitones, Hart (1981) states that "only differences of more than 3 semitones play a part in communicative situations", but smaller differences are mentioned for some tonal languages. For instance, Connell

(2000:168) reports for Mambila (northern Bantoid, with four level tones) mean differences of 11-18Hz from one level to the next, which corresponds to a mean of 2 semitones.[10] With those references in mind, we set our tentative benchmark for assessing phonological tone in CT at Kuang's 9Hz JND and 2 semitones (st): mean pitch differences greater than these two values will be considered sufficient to reflect potential tonal contrasts.

From the outset, good candidates for tone-bearing syllables are those whose accent is contrastive (not predictable from the default accented/unaccented alternation), in particular monosyllables and word-initial syllables in general. On the contrary, the open-ended default accent alternation [σ́.σ̀.σ́.σ̀.σ́...] on long words doesn't constitute a typologically common tonal pattern, but rather suggests a metrical pattern in binary feet, that is, it is more akin to stress, understood as "the linguistic manifestation of rhythmic structure" (Hayes 1995). Stress is indeed the obvious alternative to tone as an explanation for CT accent. Concretely, a prominence on every other syllable starting on the initial syllable is a widespread pattern in syllabic trochees from left to right, observable in a great number of languages such as Czech or Icelandic (Hayes 1995, Gordon 2002, Kager 2007). From the phonetic side, stress is expected to display effects of articulatory force, typically an enhanced intensity and a longer duration, as well as a fullness of articulation, among others (Fry 1955, Cutler 2005, van der Hulst 2014). Elevated pitch is also a widespread cue for stress, although it can also be explained by phrasal intonation, when intonational pitch movements are associated with prominent (stressed) syllables. In this study, we look at intensity and duration as potential cues for stress.

In what follows, we study the phonetics of accent contrasts first in monosyllables (§5), then in polysyllabic words (§6), to see what conclusions may be drawn from their prosodic properties. We present data from three native speakers, one woman (labeled F) and two men (M1 and M2), who are 20, 28 and 46 years old respectively. All of them are bilingual in Spanish. We recorded them in a quiet room with a Zoom H4n and a SM10A head-worn Shure microphone. Elicitation was conducted in Spanish by the first author and in CT by the second, who is a CT native speaker. Each sentence containing a target word was recorded between 20 and 30 times with each speaker in order to achieve statistical significance.

5. Monosyllables

We start presenting the results of our production task for monosyllables, as those are crucial for assessing the tonal hypothesis.

5.1. Overview of monosyllables

CT possesses dozens of accentual minimal pairs involving monosyllabic words, a sample of which is presented in Table 4. Note that the items in pairs (a-f) are lexically unrelated, whereas pairs (g-l) are diversely morphologically related, with the accented items being historically the basic forms. As mentioned earlier, monosyllabic verbs make up many pairs with an accented transitive 'do X' and an unaccented inchoative 'get/become Xed', like (i-m).

[10] Connell (2000) reports f0 measurements at the beginning, middle and end of the vowel. Our calculation is based on the end measurements, where f0 differences are highest.

TABLE 4

MONOSYLLABIC MINIMAL PAIRS

	Accent class 1		Accent class 2	
a.	múy	‘sapodilla (fruit sp.)’	mùy	‘go up’
b.	ʔík’	‘air, wind’	ʔìk’	‘black’
c.	níʔ	‘nose’	nìʔ	‘son-in-law’
d.	yál	‘3sg-say’	yàl	‘go down’
e.	ʔól	‘heavy’	ʔòl	‘middle’
f.	lóm	‘in vain’	lòm	‘layer (num. class.)’
g.	més	‘sweep’	mès	‘broom’
h.	ch’íx	‘thorn’	ch’ìx	‘longish thing (num. class.)’
i.	tús	‘solve’	tùs	‘become solved’
j.	nóp	‘think’	nòp	‘get used to’
k.	pás	‘do’	pàs	‘get done, to be doable’
l.	xách’	‘stretch’	xàch’	‘become stretched’
m.	nét’	‘crush’	nèt’	‘become crushed’

All consulted speakers produce the monosyllabic pairs in Table 4 as phonetically distinct in the context of carrier sentences, see below. They can also discriminate them by listening to random recordings of such words. However, we have not conducted a formal listening task to support this last claim. This will be done in future work.

The prosodic contrast is illustrated with the pairs *tús* ‘solve’ – *tùs* ‘become solved’, recorded in the carrier sentences (9), and *nóp* ‘learn’ – *nòp* ‘get used to’, recorded as in (10)-(12), with a 1st, 2nd and 3rd person subject. Note that only (11) in 2nd person constitutes a perfect minimal context like (9), whereas (10) and (12) work as near-minimal contexts. A sample recording for each pair is presented in Figure 6 and 7.

(9) a. Mú=tò s–**tús** lèʔy=e.
NEG=yet 3ERG–solve DEM=DEF
‘She doesn’t solve this one yet.’

b. Mú=tò s–**tùs** lèʔy=e.
NEG=yet IPFV–become.solved DEM=DEF
‘This one isn’t being solved yet.’

(10)a. Mú=tò j–**nóp** lìʔ=e.
NEG=yet 1ERG–think here=DEF
‘I don’t think about it here yet.’

b. Mú=tò x–ì–**nòp** lìʔ=e.
NEG=yet IPFV–1ABS–get.used here=DEF
‘I don’t get used to here yet.’

(11)a. Mú=tò x–à–**nóp** lìʔ=e.
NEG=yet IPFV–2ERG–think here=DEF
‘You don’t think about it here yet.’

b. Mú=tò x–à–**nòp** lìʔ=e.
NEG=yet IPFV–2ABS–get.used here=DEF
‘You don’t get used to here yet.’

(12)a. Mú=tò s–**nóp** lìʔ=e.
NEG=yet 3ERG–think here=DEF
‘She doesn’t think about it here yet.’

b. Mú=tò x–**nòp** lìʔ=e.
NEG=yet IPFV–get.used here=DEF
‘She doesn’t get used to here yet.’

Figure 6: Waveform, pitch track and vowel duration from the production of (9a) ‘she doesn’t solve this one yet’ (left) and (9b) ‘this one isn’t being solved yet’ (right) by speaker M1, contrasting monosyllabic verb forms

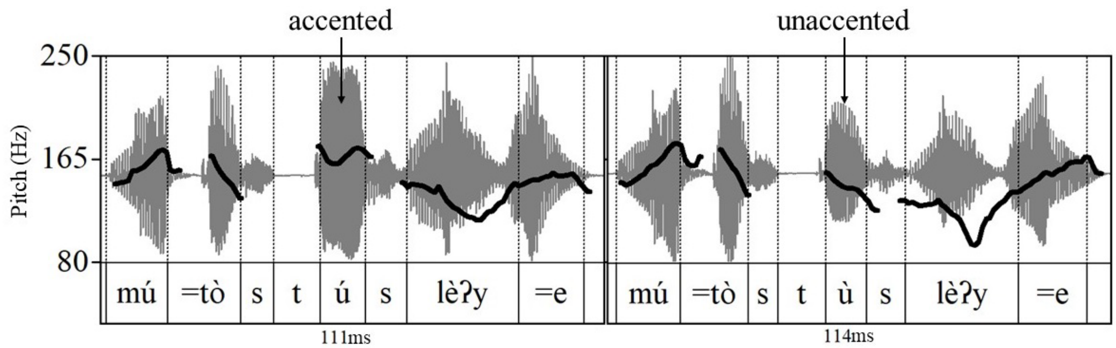
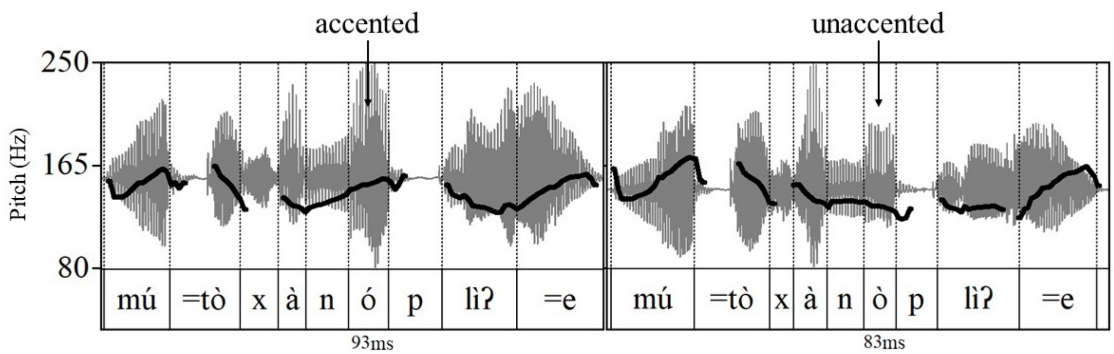


Figure 7: Waveform, pitch track and vowel duration from the production of (11a) ‘you don’t think about it here yet’ (left) and (11b) ‘you don’t get used to here yet’ (right) by speaker M1, contrasting monosyllabic verb forms



The target vowels were manually segmented using standard criteria. Pitch, intensity and duration over the target vowel were measured in the context of (9)-(12) using Praat (Boersma and Weenink 2020) and a custom script. Pitch and intensity were also measured for the vowel of the sentence-initial negation *mú* to serve as a comparison point, in order to control for the fact that the overall pitch and volume can vary from one utterance to another.

5.2. Pitch

The most salient and systematic acoustic difference between these accented and unaccented monosyllables is pitch, as measured by f0 values and semitones (st). In Table 5, we provide four ways to characterize f0. The first one is straightforward: the mean f0 of the whole target vowel. The second one is the mean f0 at 75% of the target vowel, which we claim gives a better measure for assessing tones in CT because the intended pitch is often reached late in the vowel, a phenomenon known as "peak delay" (Yip 2002: 8 and references therein; Pittayaporn 2018): as it takes time to accommodate the vocal folds to modulate pitch, the tonal target can be reached late in the vowel, producing potential rising and falling transitions even for level tones (see Figure 6 and 7 for an illustration). Those rises and falls are made explicit through the third value displayed in Table 5, which correspond to the f0 variation as the difference between the f0 values at 25% and 75% of the duration of the target vowel (positive values mean a rise and negative ones a fall). Last, the relative f0 shows the difference between the mean f0 of the target vowel and the mean f0 of the vowel of the negation *mú* ($f0_{\text{target vowel}}$ - $f0_{mú}$), which makes sure that the overall pitch level of the utterance does not distort our results (namely, the difference between both monosyllables in terms of mean relative f0 must be of the same order than the difference between their mean f0). P-values for statistical significance were obtained through a Welch's unequal variances t-test. Standard deviation is indicated in parentheses in this and all the following tables.

TABLE 5

F0 OF MONOSYLLABLE VOWELS (IN HZ AND ST): *TÚS* 'SOLVE' VS *TÙS* 'BECOME SOLVED' AND *NÓP* 'THINK' VS *NÒP* 'GET USED TO'

	N	Mean f0	Mean f0 at 75%	F0 variation	Mean relative f0
F					
tús	25	229 (9)	233 (9)	+9 (6)	-16 (8)
tùs	28	208 (5)	206 (6)	-5 (4)	-33 (4)
Difference		21 [1.7st]	28 [**2.1st**]		17
Significance		p<.0001	p<.0001	p<.0001	p<.0001
nóp	56	225 (6)	229 (7)	+9 (3)	-24 (6)
nòp	45	203 (4)	201 (4)	-4 (2)	-47 (7)
Difference		22 [1.8st]	28 [**2.3st**]		23
Significance		p<.0001	p<.0001	p<.0001	p<.0001
M1					
tús	30	168 (5)	172 (6)	+8 (2)	-4 (5)
tùs	30	142 (4)	140 (4)	-7 (2)	-33 (5)
Difference		26 [2.9st]	32 [**3.6st**]		29
Significance		p<.0001	p<.0001	p<.0001	p<.0001
nóp	90	153 (4)	156 (4)	+7 (2)	-8 (5)
nòp	90	133 (3)	132 (3)	-3 (1)	-30 (5)
Difference		20 [2.4st]	24 [**2.9st**]		22
Significance		p<.0001	p<.0001	p<.0001	p<.0001

M2					
tús	35	149 (5)	152 (5)	+6 (3)	1 (4)
tùs	26	123 (4)	120 (3)	-4 (2)	-34 (4)
Difference		26 [3.3st]	32 [**4.1st**]		35
Significance		p<.0001	p<.0001	p<.0001	p<.0001
nóp	65	139 (7)	142 (8)	+7 (3)	-9 (7)
nòp	64	115 (7)	113 (7)	-2 (2)	-36 (5)
Difference		24 [3.3st]	29 [**4st**]		27
Significance		p<.0001	p<.0001	p<.0001	p<.0001

As Table 5 shows, there is a consistent pitch difference between accented and unaccented monosyllables: the former are pronounced about 25~30Hz above the latter at 75% of the vowel (a little less over the whole vowel: 20~25Hz). This was found to be statistically significant (p<.0001) in all cases (these results remain significant after accounting for multiple measures using the Holm-Bonferroni stepwise procedure). Once normalized in semitones, the pitch contrasts at 75% exceed our 2st benchmark in all cases and reach an average of 3.2st over this data set. This evidence thus supports a phonological H vs L tone distinction. In addition, accented monosyllables show a rise of ~8Hz, while a pitch fall of ~4Hz is observable for unaccented ones. This is a common concomitant manifestation of tone in many languages, where H vs L level tones typically show a slight rising vs falling contour respectively (e.g. in San Pablo Güilá Zapotec: Arellanes Arellanes 2009). This is an effect of the above-mentioned peak delay.

5.3. Intensity

For intensity, we also distinguished the absolute peak intensity over the target vowel from the relative one, calculated by subtracting the peak intensity of the negation *mú* from the previous value, to control for the overall intensity level of the utterance. Table 6 summarizes the results.

TABLE 6

INTENSITY OF MONOSYLLABLE VOWELS (IN DB): *TÚS* 'SOLVE' VS *TÙS* 'BECOME SOLVED' AND *NÓP* 'THINK' VS *NÒP* 'GET USED TO'

	N	Absolute peak intensity	Relative peak intensity
F			
tús	25	64.6 (1.5)	1.5 (1.7)
tùs	28	63.1 (1.8)	0.1 (1.6)
Difference		1.5	1.4
Significance		p<.002	p<.004
nóp	56	60. 7 (1.3)	-1.7 (1.4)
nòp	45	57.4 (1.4)	-5.2 (2)
Difference		3.3	3.5
Significance		p<.0001	p<.0001

M1			
tús	30	68.8 (1.4)	3.9 (1)
tùs	30	64.5 (.6)	-1.5 (1)
Difference		4.3	5.4
Significance		p<.0001	p<.0001
nóp	90	62.8 (1)	1.6 (1.3)
nòp	90	58.5 (1.1)	-3.7 (1.6)
Difference		4.3	5.3
Significance		p<.0001	p<.0001
M2			
tús	35	65.9 (1.2)	2.7 (1.7)
tùs	26	62.2 (1.2)	-2.7 (1.1)
Difference		3.7	5.4
Significance		p<.0001	p<.0001
nóp	65	60.3 (2.2)	1.6 (2.7)
nòp	64	54.1 (2.2)	-5.8 (2)
Difference		6.2	7.4
Significance		p<.0001	p<.0001

Intensity appears as a statistically significant cue for accent in the case of all three speakers, as these accented syllables were generally louder than unaccented ones, with a mean difference ranging here from 1.5 to 7dB.

5.4. Duration

The results for mean duration are shown in Table 7.

TABLE 7

DURATION OF MONOSYLLABLE VOWELS (IN DB): *TÚS* 'SOLVE' VS *TÙS* 'BECOME SOLVED' AND *NÓP* 'THINK' VS *NÒP* 'GET USED TO'

	F N	Mean duration	M1 N	Mean duration	M2 N	Mean duration
tús	25	112 (16)	30	112 (9)	35	108 (11)
tùs	28	95 (10)	30	115 (7)	26	103 (6)
Difference		17		3		5
Significance		p<.0001		n.s.		p<.03
nóp	56	88 (9)	90	93 (8)	65	86 (8)
nòp	45	78 (9)	90	84 (6)	64	85 (8)
Difference		10		9		1
Significance		p<.0001		p<.0001		n.s.

Accented syllables showed a significantly longer duration in four out of six cases, but to a small extent, ranging from 5 to 17ms; in the most extreme case (case of *tús/tùs* for F), this amounts to a 6:5 ratio between accented and unaccented syllable. Besides the case of *tús/tùs* for F, all the other differences are close to or below the Just Noticeable Difference for duration, set by Stevens

(2000:228) at √(duration) for durations below 100ms. Some inter-speaker variation is noticeable, with speaker F making the greatest use of duration differences and M2 the least. In light of this, duration appears to be a weaker and less systematic cue for accent than pitch and intensity with this kind of monosyllabic word.

5.5. Tonal analysis

The quantitative evidence provided so far (concerning only monosyllables) is compatible with an analysis in terms of phonological tone, namely a H/L tone contrast. The pitch difference of 25~30Hz and ~3st between accented and unaccented monosyllables corresponds to that registered between H and L tones of languages with two level tones, as mentioned earlier. That difference is well above the 9Hz "just noticeable difference" (JND) for lexical tone (Kuang 2013: 3) and above the threshold of 2st we established earlier. The fact that accented and unaccented syllables tend to display a small rising vs falling contour respectively is also a common behavior observed in level tones.[11]

As for the fact that intonation and duration correlate with accent, this also can be explained as a phonetic effect of tone (Zee 1978, Whalen and Xu 1992, Lian-Hee 2019: 15). For instance, Zee (1978) shows that in Taiwanese, a language with L, M (mid) and H tones, H shows a higher overall intensity than M, and M a higher overall intensity than L, with differences ranging approximately from 2 to 4dB depending on the speaker. Regarding duration, H and M are longer than L roughly to a 5:4 ratio (H is also slightly longer than M but not significantly so).[12] Although there are nuances in the values, similar contrasts were shown in CT (compared to Taiwanese, the contrast in CT appears smaller for duration and greater for intensity, the latter at least for speaker M2). With this in mind, the acoustic results presented above are consistent with a tonal analysis of accent for monosyllables.

A tonal analysis in CT is also promising for interpreting some of the prosodic phenomena reviewed in §3, Table 3. On the one hand, deaccenting (§3.2), whereby accented syllables get deaccented before another accented syllable, fits a tonal approach well: certain tones are notoriously subject to dissimilation effects, whereby adjacent identical tones are prohibited. This is captured by the Obligatory Contour Principle (Odden 1986), in this case disallowing adjacent H tones –OCP(H)– which amounts to a simple deletion rule in CT: "delete a H before an adjacent H in the same IP".

On the other hand, the effect of phrasal intonation also lends itself well to a tonal account. As discussed in §3.4, the last syllable of IPs is used for phrasal intonation, displaying the contour required by the sentence pragmatics and the speaker's attitude and emotion. Concretely regarding monosyllabic minimal pairs, it means that they are neutralized in that position, as shown in (13), in contrast to (9) above (following the convention introduced earlier, cf. fn. 5, the IP-final syllable is not marked for accent). No matter what intonation is given to the sentence in (13), there is simply

[11] Other factors may influence pitch contours beyond accent. In particular, there is often a pitch drop before glottalized consonants /b p' t' k' ts' ch'/ and the glottal stop /ʔ/, independently of accent. This effect is linked to the laryngealization of vowels adjacent to those consonants, as has been documented in several Mayan languages (Bennett 2016: 484 and references therein; see also Bennett et al. 2022b for the same phenomenon in Uspanteko), including CT: Herrera (2013, 2014). This is visible in Figure 3 (§3.2) with the minimal pair *sníʔ* 'her nose' vs *snìʔ* 'her son-in-law': both vowels /i/ show a pitch drop before /ʔ/. Nevertheless, this does not prevent this pair from being contrastive. We leave a phonetic account of the interaction between accent and laryngealized voicing for future research.

[12] Faytak and Yu (2011) argue that the converse is typologically more pervasive –the lower the tone, the longer the duration– but that exceptions are not rare, in particular due to interference from stress.

no way to prosodically discern both meanings. The same happens with words uttered in isolation (citation forms), as they constitute IPs by themselves.

(13) Mú=tò s–**tus**.
NEG=yet 3ERG/IPFV–solve/become.solved
'She doesn't solved it yet.' or 'It isn't being solved yet.'

Herrera (2013, 2014) studied words in isolation, this explains why she did not find evidence for tone (or accent) contrasts. In the framework of autosegmental-metrical theory (Pierrehumbert and Beckman 1988, Ladd 2008), intonational events at the end of the IP are analyzed as boundary tones (T%: H% or L%, depending on the contour). With a tonal approach to CT accent, the neutralization in (13) is straightforward: a boundary tone on the last syllable of the IP overrides the underlying tone, be it the L tone of *tùs* 'become solved' or the H tone of *tús* 'solve', making both words identical. That is, there is a tone simplification rule operating in CT: the underlying tone is overriden to avoid tonal crowding on that syllable.[13] Likewise, monosyllables uttered in isolation necessarily receive a boundary tone that neutralizes the contrast between accent classes, because they must constitute an IP on their own.

The specific overriding of H tones on the IP-final syllable gains support from a connected phenomenon: no deaccenting obtains from the last syllable of the IP onto the penultimate syllable. Deaccenting was illustrated in §3.2 by example (4), repeated here for convenience, where the verb *tsák* 'grab' gets deaccented before *níʔ* 'nose', but not before *nìʔ* 'son-in-law'. We show in (14) the same examples modified by omitting the final definite clitic =*e*, which is optional with possessed nouns (López Mendoza 2022),[14] leaving the object in the IP-final position.

(4)a. Là s–ts**à**k s–n**í**ʔ=e.
PFV 3ERG–grab 3POSS–nose=DEF
'She grabbed her nose.'

b. Là s–ts**á**k s–n**ì**ʔ=e.
PFV 3ERG–grab 3POSS–son-in-law=DEF
'She grabbed her son-in-law.'

(14) Là s–ts**á**k s–n**i**ʔ.
PFV 3ERG–grab 3POSS–nose/son-in-law
'She grabbed her nose.' or 'She grabbed her son-in-law.'

As expected, the minimal pair is then neutralized, just like in (13) above. More surprisingly, the preceding verb here remains accented, even if the object is interpreted as 'nose' (and no matter whether the boundary tone is H% or L%). We suggested above that OCP(H) explains

[13] According to Hyman and Monaka (2011), this corresponds to the "submission" strategy, whereby "the intonational tones invade and override the lexical tones". Similar phenomena, which potentially lead to tonal neutralization, have been documented in other Mesoamerican languages, such as Uspanteko (Bennett et al. 2019, 2022a, 2022b), Mazahua (Pike 1951) and Acazulco Otomi (Hernández Green 2015:28). Van der Hulst (2011:1005) also discusses the absence of tonal contrasts on the final syllable as a way "to leave room for intonational tones" in some languages.
[14] See Aissen (2017) on the prosodic behavior of this clitic =*e* in Zinacantec Tsotsil.

deaccenting. Its lack of application in (14) can be accounted for by the fact that the final syllable bears a boundary tone, H% or L%, instead of H: lexical tones and boundary tones are different kinds of phonological objects. We therefore propose that the OCP targets here the former but not the latter.

In summary, the evidence discussed in this section supports the hypothesis of a H/L tonal opposition for analyzing CT accent. First, fundamental frequency stands out as one of the key ingredients of the acoustic contrast between monosyllables. Intensity, and to a lesser extent duration, also appear as cues to accent, but they can be explained as a side effect of tone. Finally, two phenomena which are part of the larger equation of CT prosody – deaccenting and IP-final effects– can be adequately analyzed through tonal mechanisms involving H tones. We take up this discussion in §7.

6. Polysyllables

We now move on to explore the phonetics of accent in a few selected polysyllables to assess pitch distinctions that could count as evidence for tone, as well as correlates of stress, namely intensity and duration.

6.1. Overview of polysyllables

We present a sample of polysyllabic words from both accent classes in Table 8, including some minimal pairs.

TABLE 8
POLYSYLLABLES

σ	Accent class 1		Accent class 2	
2	ʔólìl	'child'	ʔòlíl	'half'
	ts'ákàl	'complete'	ts'àkál	'afterward'
	túsùl	'sticking out'	tùsúl	'green bean'
	pótò(j)	'guava'	pòkó	'old'
3	túsòbíl	'solution'	tùsòbíl	'comb'
	ʔínìtám	'iguana'	bànkìlál	'big brother'
	lóʔbàjél	'fruit'	màtànál	'gift'
4	(j)póxtàvánèj	'doctor'	tòjbàlálìl	'slave'
	ʔám̱tèjébàl	'tool'	ʔànìlájàn	'you run (irrealis)'
5	(j)chánùm̱tésvànéj	'teacher'	nàtùm̱tésvàném	'it has made people taller'
	xchíkìltáòjún	'she has tickled me'	svàlàk'pátìój	'she has turned her back on him'
6	chánùm̱tésvànémòt	'you have taught people'	nàtùm̱tésvànémòt	'you have made people taller'
	xchíkìltáòjúnìk	'they have tickled me'	svàlàk'pátìójùn	'she has turned her back on me'

As mentioned earlier, accent patterns for polysyllables contrast in the initial syllable, plus the second one for disyllables. From the third syllable onwards, the same binary alternation is observed, which brings to mind metrical stress, specifically a trochaic pattern, more than tone. Both perspectives –tone and stress– can be reconciled if the tonal specification concerns the initial syllable, while metrical stress explains the rest; this is the analysis we develop in §7.

6.2. Results

We conducted a production study with the near-minimal pair *pótò* (class 1) 'guava' vs *pòkó* (class 2) 'old', already mentioned in §3.4 above. This pair has the advantage of displaying a symmetric syllabic pattern CV.CV with the same vowel and voiceless plain plosives, hence making a good starting point to compare accented and unaccented syllables in a disyllabic word. Note that speakers M1 and M2 say *pótò**j*** with a final velar fricative instead of *pótò*, affecting to some degree the second vowel, possibly shortening it. Nevertheless, in CT this velar fricative has a quite weak pronunciation in coda, to the point that it is not always easily audible, so its impact might be modest anyway.[15] This pair was recorded in the carrier sentence already seen in (8) above; see Fig. 5 (right) for a sample. The results are presented in Table 9 (standard deviation below each value in parentheses). Differences V2-V1 are added for comparison.

TABLE 9

F0 (in Hz and st), intensity (in dB) and duration (in ms) of *pótò(j)* 'guava' and *pòkó* 'old' in carrier sentence of (8)

		Mean f0			Mean f0 at 75%			F0 variation		Peak intensity			Duration		
	N	V1	V2	V2-V1	V1	V2	V2-V1	V1	V2	V1	V2	V2-V1	V1	V2	V2-V1
F															
[pó.tò]	30	241	230	-11	249	222	**-27**	+14	-16	66.5	64.0	**-2.5**	113	85	**-28**
		(3)	(4)	[0.8st]	(5)	(5)	**[2st]**	(4)	(4)	(1.7)	(1.2)		(9)	(10)	
[pò.kó]	30	223	245	+22	221	248	**+27**	-3	+6	62.6	67.1	**+4.5**	85	134	**+49**
		(6)	(7)	[1.6st]	(6)	(8)	**[2st]**	(2)	(4)	(1.7)	(1.5)		(10)	(12)	
M1															
[pó.tòj]	30	162	150	-12	166	141	**-25**	+7	-14	66.5	62.7	**-3.8**	86	65	**-21**
		(6)	(3)	[1.3st]	(6)	(3)	**[2.8st]**	(2)	(3)	(1.2)	(1.1)		(7)	(13)	
[pò.kó]	30	142	164	+22	139	168	**+29**	-6	+10	61.8	66.2	**+4.4**	63	97	**+34**
		(4)	(4)	[2.5st]	(4)	(5)	**[3.3st]**	(2)	(3)	(1.5)	(1.1)		(7)	(9)	
M2															
[pó.tòj]	30	162	141	-21	169	133	**-36**	+12	-8	68.7	64.1	**-4.6**	81	56	**-25**
		(8)	(7)	[2.4st]	(9)	(7)	**[4.1st]**	(3)	(5)	(1.4)	(2.1)		(9)	(11)	
[pò.kó]	30	134	167	+33	131	171	**+40**	-5	+8	61.9	68.4	**+6.5**	59	113	**+54**
		(5)	(7)	[3.8st]	(5)	(8)	**[4.6st]**	(3)	(4)	(2.1)	(1.2)		(10)	(11)	

These results confirm the contrastive prosodic patterns described as σ́.σ̀ (pattern I) vs σ̀.σ́ (pattern II). They also match accent manifestation in monosyllables (cf. §5.2-5.4): accented syllables are more prominent than unaccented ones trough their greater pitch, intensity and duration. Regarding pitch, the mean f0 differences at 75% of the vowel range from 25 to 40Hz and from 2 to 4.6st, that is, they are in the expected range of values for contrastive level tones, crucially above 9Hz/2st (cf. §3.5).[16]

Duration in these data turns out to be more important as a cue to accent than with monosyllables. As commented in §5.4, duration is a non-systematic cue to accent in monosyllables, with measured contrasts ranging from none to a maximal 6:5 ratio between accented and non-accented syllables. In Table 9, the disparity is more pronounced: the approximate duration ratio between both types of syllables ranges here from 4:3 (*pótòj*, F and M1) to almost 2:1 (*pòkó*, M2). Our impression is that this is indeed a quite systematic trend for polysyllables, in

[15] The lack of consequence of this final fricative for our purposes is confirmed by the fact that the values obtained from speaker F coincide with those from M1 and M2, as shown in Table 9.

[16] The mean f0 over the whole vowel yields some values (for speaker F and M1 with *pótòj*) just above 9Hz and below the adopted 2st threshold. We argued above that the f0 at 75% is a better measure of the pitch target than the overall mean f0, as it takes into account the non-negligible effect of the divergent f0 contours caused by the phenomenon of peak delay.

which unaccented syllables tend to shorten in contact with accented syllables. This is expected with stress, as stress concerns the relative prominence between syllables of the same word (cf. Hyman 2006, stress is syntagmatic, whereas tone is paradigmatic). This hints at the relevance of stress in CT.

We have insisted on the fact that the accent contrast between polysyllabic words is condensed on the first syllable. To test this observation, we recorded the same pair of words in isolation. As shown above, in this context the final syllable bears a pitch excursion associated to IPs, canceling accent contrasts on that syllable. Table 10 presents the results, see Figure 5 (upper- and lower-left) above for a sample.

TABLE 10

F0 (IN HZ AND ST), INTENSITY (IN DB) AND DURATION (IN MS) OF *PÓTÒ(J)* 'GUAVA' AND *PÒKÓ* 'OLD' IN ISOLATION

		Mean f0			Mean f0 at 75%			F0 variation		Peak intensity			Duration		
	N	V1	V2	V2-V1	V1	V2	V2-V1	V1	V2	V1	V2	V2-V1	V1	V2	V2-V1
F															
[pó.tò]	30	237	243	**+6**	243	243	**0**	+10	-1	65.5	65.1	**-0.4**	151	170	**+19**
		(4)	(7)	[0.4st]	(4)	(9)	[0st]	(3)	(6)	(1.1)	(1.2)		(10)	(12)	
[pò.kó]	30	220	250	**+30**	219	265	**+46**	-2	+27	62.5	65.9	**+3.4**	89	189	**+100**
		(5)	(5)	[2.2st]	(7)	(7)	[3.3st]	(4)	(7)	(1.6)	(1.1)		(14)	(11)	
M1															
[pó.tòj]	30	154	168	**+14**	158	167	**+9**	+7	-1	65.0	64.8	**-0.2**	95	150	**+55**
		(4)	(6)	[1.5st]	(4)	(6)	[1st]	(3)	(5)	(0.9)	(1)		(9)	(16)	
[pò.kó]	30	141	168	**+27**	138	173	**+35**	-5	+12	60.6	66.2	**+5.6**	67	157	**+90**
		(5)	(3)	[3st]	(4)	(5)	[3.9st]	(3)	(5)	(1.5)	(1.6)		(8)	(15)	
M2															
[pó.tòj]	30	143	158	**+15**	146	158	**+12**	+4	+1	63.5	65.8	**+2.3**	77	197	**+120**
		(4)	(8)	[1.7st]	(5)	(8)	[1.4st]	(4)	(3)	(1.2)	(2.1)		(6)	(22)	
[pò.kó]	30	124	160	**+36**	121	162	**+41**	-6	+6	61.0	65.9	**+4.9**	61	189	**+128**
		(6)	(9)	[4.4st]	(6)	(9)	[5.1st]	(4)	(4)	(2.1)	(2.4)		(10)	(26)	

Observe that V2 is more or less uniformly prominent here, which is expected from its IP-final position. On the other hand, the prosodic contrast between V1 and V2 is more important for *pòkó* 'old' than it is for *pótò(j)* 'guava': just as in the previous carrier sentence context, V1 in *pótò(j)* displays greater pitch, intensity and duration values than in *pòkó* 'old', and its pitch shows a slight rise instead of a fall. We applied a Welch's t-test comparing the differences V2-V1 for [pó.tò(j)] and [pò.kó] and found a high degree of significance ($p<.0001$, confirmed through the Holm-Bonferroni stepwise procedure) for all parameters but one –the duration contrast for M2 (V2-V1=120 vs 128ms)–, which recalls the inter-speaker variation in the extent duration is used as a cue to accent in monosyllables.[17] Overall, these results confirm that the load of the contrast between disyllables can be assumed by the sole first syllable.

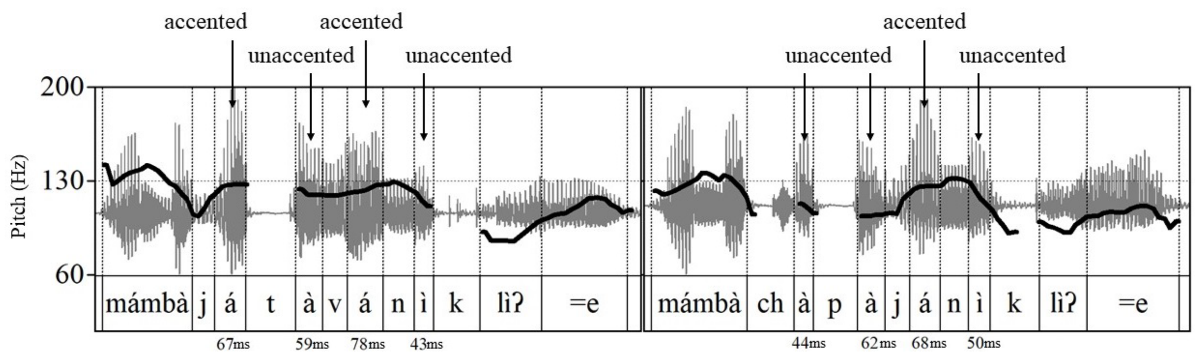
[17] Note that M2 is the speaker for whom duration was least significant in the contrast between monosyllables, cf. §5.4. This could explain why this same speaker does not rely on duration in disyllables when the contrast is limited to the first syllable.

Up to this point, the evidence validates the contrastiveness of accent patterns for disyllables and the fact that the contrast still operates when confined to the sole first syllable. This is coherent with the hypothesis that the two possible accent patterns basically stem from a H vs L tonal opposition on the initial syllable. At the same time, the importance of duration as a cue to accent suggests that stress is also relevant.

In what follows, we continue this exploration of polysyllables by studying the pair of examples in (15), featuring four-syllable inflected forms of the verbs *játàv* 'run away' (class 1) vs *chàpáj* 'get organized' (class 2), selected for having three comparable /Ca/ syllables. See a sample in Figure 8.

(15)a.	Mámbà	játàv–án–ìk	lìʔ=e.
	NEG	run.away–2ABS.IRR–PL	here=DEF
	'Don't run away from here.'		
b.	Mámbà	chàpàj–án–ìk	lìʔ=e.
	NEG	get.organized–2ABS.IRR–PL	here=DEF
	'Don't get organized here.'		

Figure 8: Waveform, pitch track and vowel duration from the production of (15a) 'don't run away from here' (left) and (15b) 'don't get organized here' (right) by speaker M2, showing the contrasting prosodic pattern of class 1 and 2 tetrasyllables



We comment first on the f0 measures presented in Table 11. Intensity and duration are presented together in Table 13 below.

TABLE 11

F0 of V1/2/3/4 for *játàvánìk* 'you (pl.) run away (irrealis)' and 'you (pl.) get organized (irrealis)' (in Hz)

		Mean f0				Mean f0 at 75%				F0 variation			
	N	V1	V2	V3	V4	V1	V2	V3	V4	V1	V2	V3	V4
F													
[já.tà.vá.nìk]	20	220	208	213	203	220	205	213	202	0	-7	0	-3
		(5)	(6)	(5)	(6)	(5)	(6)	(5)	(6)	(3)	(2)	(2)	(1)
[chà.pà.já.nìk]	20	211	197	215	208	205	195	216	206	-12	-3	+1	-4
		(5)	(5)	(6)	(7)	(5)	(5)	(7)	(8)	(2)	(3)	(2)	(2)
M1													
[já.tà.vá.nìk]	20	143	137	141	137	144	134	142	135	+2	-8	+2	-5
		(5)	(4)	(3)	(5)	(4)	(4)	(4)	(5)	(3)	(2)	(1)	(1)
[chà.pà.já.nìk]	20	139	134	145	145	136	130	147	143	-7	-7	+3	-3
		(4)	(4)	(3)	(4)	(3)	(5)	(4)	(4)	(3)	(3)	(2)	(1)
M2													
[já.tà.vá.nìk]	20	122	113	116	105	122	111	117	102	+1	-5	+2	-6
		(8)	(8)	(6)	(5)	(8)	(7)	(6)	(4)	(1)	(6)	(1)	(2)
[chà.pà.já.nìk]	20	107	105	115	110	105	103	117	107	-4	-6	+4	-6
		(5)	(5)	(5)	(5)	(4)	(5)	(5)	(5)	(2)	(5)	(2)	(2)

Pitch differences are modest between adjacent vowels, even taking into account the value at 75%: as summarized in Table 12 below, far from the 25Hz difference or more observed among monosyllables, the difference is less than 20Hz, and in some cases it is smaller than the 9Hz JND. (The V1/V2 difference in [chà.pà.já.nìk], not expected to be significant as it concerns two unaccented syllables, appears in a smaller size in Table 12).

TABLE 12

Pitch differences at 75% from data of Table 11 (in Hz and st)

		V1/V2	V2/V3	V3/V4
F	[já.tà.vá.nìk]	-15 [1.2st]	+8 [0.7st]	-11 [0.9st]
	[chà.pà.já.nìk]	-10 [0.9st]	+21 [1.8st]	-10 [0.8st]
M1	[já.tà.vá.nìk]	-10 [1.2st]	+8 [1st]	-7 [0.9st]
	[chà.pà.já.nìk]	-6 [0.8st]	+17 [2.1st]	-4 [0.5st]
M2	[já.tà.vá.nìk]	-11 [1.6st]	+6 [0.9st]	-15 [2.4st]
	[chà.pà.já.nìk]	-2 [0.3st]	+14 [2.2st]	-10 [1.5st]

However, pitch variation over the word correlates well with accent changes: f0 always drops from an accented to a unaccented vowel and conversely rises from an unaccented to an accented vowel. On average over this limited corpus and pooling pitch values at 75% from all three speakers, pitch drops 10.3Hz/1.2st from accented to unaccented syllables and rises 12.3Hz/1.4st in the opposite situation, which is slightly above the 9Hz JND, yet below our proposed 2st threshold. At the same time, on each syllable pitch varies in the way already noted before: it tends to rise (or at least, not to fall) on accented syllables (on average: +1.7Hz) and to drop on unaccented ones (on average: -5.9Hz), cf. Table 11, rightmost columns. Therefore, despite the small differences, pitch here correlates with accent in a systematic way, which still leaves open the possibility of a tonal approach insofar as additional factors can be invoked that mitigate pitch differences. We return to this point below, after reviewing intensity and duration measurements.

TABLE 13

PEAK INTENSITY (IN DB) AND DURATION (IN MS) OF V1/2/3/4 FOR *JÁTÀVÁNÌK* 'YOU (PL.) RUN AWAY (IRREALIS)' AND 'YOU (PL.) GET ORGANIZED (IRREALIS)'

		Peak intensity				Duration			
	N	V1	V2	V3	V4	V1	V2	V3	V4
F									
[já.tà.vá.nìk]	20	65.4	64.0	65.0	53.6	71	60	91	33
		(1.5)	(1.6)	(1.5)	(1.3)	(7)	(8)	(9)	(4)
[chà.pà.já.nìk]	20	60.5	61.9	60.8	54.5	48	47	87	37
		(1.5)	(1.4)	(1.9)	(1.5)	(7)	(7)	(8)	(5)
M1									
[já.tà.vá.nìk]	20	64.3	64.0	62.7	55.7	68	57	80	37
		(1.1)	(1.1)	(1.4)	(0.8)	(9)	(6)	(5)	(7)
[chà.pà.já.nìk]	20	61.7	60.8	61.6	57.0	43	52	72	42
		(1.7)	(1.2)	(0.9)	(0.9)	(6)	(8)	(5)	(5)
M2									
[já.tà.vá.nìk]	20	61.6	58.8	60.0	53.1	63	49	77	36
		(2.8)	(2.7)	(1.8)	(1.7)	(7)	(7)	(11)	(6)
[chà.pà.já.nìk]	20	56.0	54.2	56.6	54.2	44	43	80	47
		(2)	(2.9)	(1.4)	(1.4)	(4)	(6)	(7)	(4)

For intensity and duration, V4 is not meant to be comparable here because of its different vowel quality ([i] instead of [a]) and the fact it is in a closed instead of open syllable. Focusing on the first three vowels, duration appears as the most systematic cue to accent. Where the contrast is the smallest, namely between V1 and V2 in [**já.tà**.vá.nìk], with a difference of respectively 11ms (F and M1) and 14ms (M2), it is statistically significant in all three cases (F: p<.0002; M1 and M2: p<.0001); other contrasts are even greater.

On the contrary, the difference in intensity between the syllables in comparison does not appear to be consistently significant across speakers. In [já.tà.vá.nìk], only between V1 and V2 and for two out three speakers is there a significant intensity contrast (F: p<.008, M2: p<.02; n.s. for M1: p>.28), whereas in all cases V3 fails to achieve significance with respect to V2 (and is even inferior to V2 for M1). Note also that in [chà.pà.<u>já</u>.nìk] the accented V3 has no intensity prominence: for M1 and M2 it is virtually in a tie with V1, and for F it is even inferior to V2, despite the fact V2 has a low prosodic profile in terms of both pitch and duration. This is a sign of a lack of systematic correlation between accent and intensity here.

In sum, this pair of four-syllable words does confirm the described accent patterns σ́.σ̀.σ́.σ̀ (pattern I) vs σ̀.σ̀.σ́.σ̀ (pattern II) and shows that accent in long words is cued primarily by pitch and duration, and only unsystematically by intensity. Pitch differences here are conspicuously smaller than what we registered for monosyllables, with values closer to 10 Hz than to 25Hz, and even inferior to the 9Hz JND in some cases. In part, this could be a side-effect of the well-known phenomenon of polysyllabic shortening, whereby the duration of vowels in long words tends to decrease. For one thing, less duration means less time to modulate the f0, thus compressing the overall pitch range. But it is probably also a consequence of the lack of contrastiveness of accent beyond the first syllable: as a predictable pattern, speakers can afford to under-intonate it to some extent, as long as the two contrasting patterns can be distinguished. A similar situation is highlighted for Uspanteko by Bennett et al. (2022a), who report f0 differences of ~10Hz between H-toned and toneless syllables: they link this small value to the low functional load of tone in that language[18] (see also Hall et al. 2018).

The importance of duration as a cue to accent in polysyllables in general, both for disyllables and for longer words, advocates for analyzing these prosodic patterns in terms of stress. This is coherent with the view that the default alternation between accented and unaccented syllables most likely manifests a metrical pattern in syllabic trochees.

6.3. Main stress

If stress is involved, then an important question is whether words with multiple accented syllables show evidence for a main stress—that is, whether stress in CT fulfils culminativity ("at most one syllable with the highest degree of prominence"), which is generally considered a part of the definition of stress (van der Hulst 2014). However, this question proves difficult to answer, because there is often no marked prosodic asymmetry between different accented syllables. If such an asymmetry does surface, the general tendency is that the leftmost accented syllable (first syllable in class 1 words, second or third syllable in class 2 words) is the most prominent, at least in terms of pitch. This is the case for instance in [já.tà.vá.nìk], cf. Table 11, though it is significantly so only for F, $p<.0001$, and for M2, $p<.03$, not for M1.[19] The same is also visible in the sample words shown earlier in §3.1, see Figure 1. For instance in *spójtànójtàlél* (Figure 1 right) there is a stepwise lowering in the mean f0 of the three accented syllables, respectively 165, 153 and 147Hz.

However, this f0 lowering could also be an effect of declination, which is the typical downward trend of the pitch over IPs (Cohen and 't Hart 1967, Bolinger 1978, Pierrehumbert 1980, Ladd 2008), which would imply that the initial higher pitch does not necessarily correspond to a greater metrical prominence. However, there are reasons to believe that the leftmost accented syllable carries the main stress: under emphasis, for instance in pre-verbal contrastive focus position, that syllable is the one whose prominence gets exacerbated. This confirms its special metrical status as, in that context, an intonational pitch accent can be said to dock on the word's metrically strongest syllable (cf. Gussenhoven's 1983 "Focus-to-Accent"). This is illustrated in Figure 9 with *póxtàvánèj* 'doctor' (class 1) in focus: the pitch excursion on the initial syllable is notable, elevating the mean f0 of V1 up to 202Hz (with a rise from 193 to 216Hz), whereas V3 (syllable [vá]) shows a lower mean f0 of 170Hz. With a class 2 word such as *nàtùmtéslànél* 'making them longer', the pitch accent also appears on the leftmost accented syllable, which in this case is V3. See in Figure 10 how the third syllable [tés] stands out from the rest of the syllables (mean f0 = 170Hz, compared with the subsequent accented syllable V5 [né(l)], mean f0 = 124Hz).

[18] Although in another study, Bennett et al. (2022b) find a more substantial effect of H tone on f0 in Uspanteko once intonation is better controlled.

[19] Note that V3 in [já.tà.vá.nìk] is longer than V1 for all three speakers ($p<.0001$), but the fact that V1 is followed by a voiceless stop (/t/) and V3 by a sonorant (/n/) might play a role in the relative shortness of V1. We leave this issue for further research.

Figure 9: Waveform, pitch track and vowel duration from the production of *PÓXTÀVÁNÈJ là kíl=e* ‘it’s a DOCTOR that I saw’ (speaker M1), showing the effect of focus

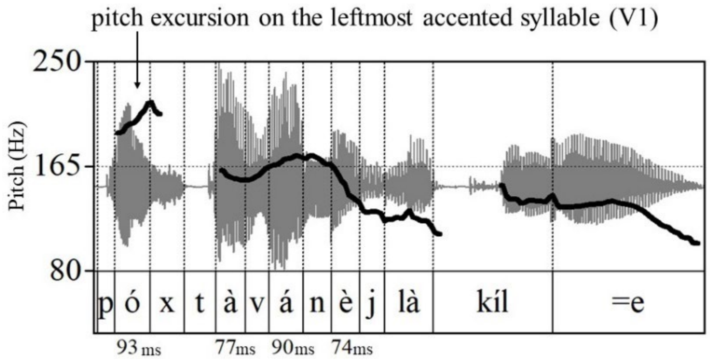
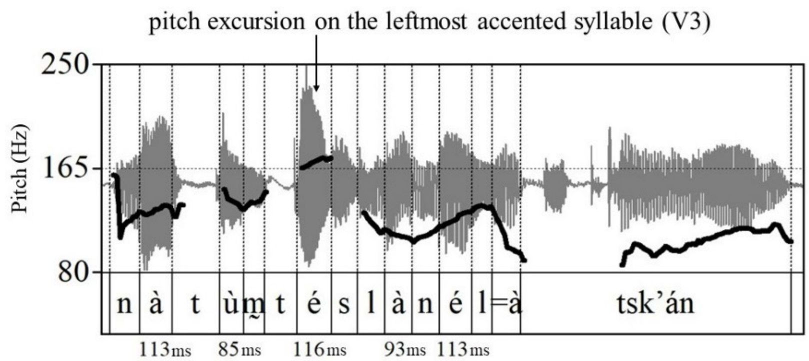


Figure 10: Waveform, pitch track and vowel duration from the production of *NÀTÙM̰TÉSLÀNÉL=(l)à tsk’án* ‘it’s MAKING THEM LONGER that is needed’ (speaker M1), showing the effect of focus



This preliminary evidence thus points to a main stress on the leftmost accented syllable. Nevertheless, we acknowledge that more research is needed to settle this issue, in particular with quantitative data, so we limit ourselves to talking about stress at the foot level in the following section.

7. Analysis

To sum up, while we argued in §5 that an analysis in terms of a H/L tonal opposition can explain the phonetics of accentual contrasts between monosyllables, this cannot be the whole story for polysyllables. For one thing, although they display systematic variations in pitch in accordance with the described accentual patterns, they also show evidence of metrical stress through a rhythmic pattern of prosodic prominence and the fact that this prominence involves duration to an extent that a purely tonal perspective cannot account for.

However, tone and stress by no means are mutually exclusive. Interactions between both types of features are indeed put forward by many authors in describing languages as diverse as Uspanteko (Bennett and Henderson 2013; Bennett et al. 2022a) and Goizueta Basque (Hualde et al. 2008), see also Jun’s (2014) typology and references therein. In fact, as argued by de Lacy (1999, 2002), an affinity exists between H tones and foot heads, as well as between L tones and foot non-heads. This affinity can translate both into tone-driven stress (e.g. stress attracted to H-toned syllables, as in Tibetan, Meredith 1990; cf. also stress retraction to H-toned penults in

Uspanteko, Bennett and Henderson 2013) or into stress-driven tone. The latter is illustrated by de Lacy (2002) with Lamba (Bantu), based on Bickmore (1995), where H-toned words display a H tone on every foot head inside their stress domain, which is parsed into trochaic feet (other proposed cases of metrical tone include Zec 1999 on Neo-štokavian Serbo-Croatian and Martínez-Paricio and Kager 2016 on Chugach Alutiiq). In both cases, the stress/tone relationship can be analyzed as mutual repulsion between L tones and foot heads (*HD/L constraint) and between H tones and foot non-heads (*NON-HD/H constraint).

In what follows, we sketch an analysis in derivational terms accounting for CT accent as H/L tone and stress. Based on de Lacy's insights, we propose that CT displays both tone-driven parsing and H-insertion on underlyingly toneless foot heads. In a nutshell, in our analysis, accented syllables correspond to syllables both stressed and H-toned, while unaccented syllables cover all other cases (stressed but not H-toned, unstressed, with or without an underlying L tone).

First, the default pattern whereby accented and unaccented syllables alternate in a strictly binary fashion is analyzed as parsing into trochees from left to right, where each foot head receives a H tone, while syllables that end up toneless are interpreted as phonetically L-toned. This is illustrated in (16) by the two suffixed forms of the verb *pój* 'take away' from example (1) above, forming respectively four- and five-syllable words. The fact that the final syllable is accented in (16b) shows that CT admits degenerate monosyllabic feet; this is also the case for monosyllables, such as *sníʔ* 'her nose', (17). We take this as an indication that parsing in this language must be exhaustive (no syllable is left unfooted inside a prosodic word).

	Underlying		Parsing		H-insertion
(16)a.	/spo.joj.ta.lel/	→	(ˈspo.joj)(ˈta.lel)	→	(ˈspo .joj) (ˈta .lel) | | H H
b.	/spoj.ta.noj.ta.lel/	→	(ˈspoj.ta)(ˈnoj.ta)(ˈlel)	→	(ˈspoj .ta) (ˈnoj .ta) (ˈlel) | | | H H H
(17)	/sniʔ/	→	(ˈsniʔ)	→	(ˈsniʔ) | H

Under this proposal, class 1 words are underlyingly toneless and acquire tone only as part of the derivation; this is stress-driven tone. On the contrary, class 2 words bear a lexical L tone on their initial syllable. Positing a phonological L tone on class 2 words, while class 1 words are analyzed as lexically toneless, has the merit of reflecting the marked status of class 2. As argued in §3.1, class 2 is marked in several respects, as the continuation of proto-forms containing a /h/

segment, either as part of the lexical root or as a derivational morpheme, and also by the fact class 2 is numerically smaller than class 1.

On monosyllables, this lexical L tone prevents the metrical H tone from linking to the syllable, as *snìʔ* ‘her son-in-law’ in (18). On disyllables, such as *pòkó* ‘old’ in (19), the initial L provokes a stress shift: the foot is parsed as an iamb (i.e., with final prominence) instead of a trochee, because of the mentioned mutual repulsion between L tones and foot heads; this is tone-driven stress. The same occurs to the initial foot of longer L-toned words, such as *chàpàjánìk* ‘you pl. get organized (irrealis)’, as in (20), to the difference that the second syllable in this case surfaces as unaccented because of the OCP(H) (deaccenting), by which H tones are deleted before a H-toned syllable.

<table>
<tr><td></td><td colspan="2">Underlying</td><td></td><td colspan="2">Parsing</td><td></td><td colspan="2">H-insertion</td></tr>
<tr><td>(18)</td><td colspan="2">/sniʔ/</td><td>→</td><td colspan="2">(ˈsniʔ)</td><td>→</td><td colspan="2">(ˈsniʔ)</td></tr>
<tr><td></td><td>|</td><td></td><td></td><td>|</td><td></td><td></td><td>|</td><td>*</td></tr>
<tr><td></td><td>L</td><td></td><td></td><td>L</td><td></td><td></td><td>L</td><td>H</td></tr>
<tr><td>(19)</td><td>/po</td><td>.ko/</td><td>→</td><td>(po</td><td>.ˈko)</td><td>→</td><td>(po</td><td>.ˈko)</td></tr>
<tr><td></td><td>|</td><td></td><td></td><td>|</td><td></td><td></td><td>|</td><td>|</td></tr>
<tr><td></td><td>L</td><td></td><td></td><td>L</td><td></td><td></td><td>L</td><td>H</td></tr>
</table>

<table>
<tr><td></td><td colspan="2">Underlying</td><td></td><td colspan="2">Parsing</td><td></td><td colspan="3">H-insertion</td><td></td><td colspan="3">OCP(H)</td></tr>
<tr><td>(20)</td><td>/cha</td><td>.pa.ja.nik/</td><td>→</td><td>(cha</td><td>.ˈpa)(ˈja.nik)</td><td>→</td><td>(cha</td><td>.ˈpa)</td><td>(ˈja .nik)</td><td>→</td><td>(cha</td><td>.ˈpa)</td><td>(ˈja .nik)</td></tr>
<tr><td></td><td>|</td><td></td><td></td><td>|</td><td></td><td></td><td>|</td><td>|</td><td>|</td><td></td><td>|</td><td>ǂ</td><td>|</td></tr>
<tr><td></td><td>L</td><td></td><td></td><td>L</td><td></td><td></td><td>L</td><td>H</td><td>H</td><td></td><td>L</td><td>H</td><td>H</td></tr>
</table>

The head status of the second syllable [pa] in (20) may not be phonetically salient (cf. intensity and duration measures for this word in Table 13), but an additional phenomenon strengthens the case for a stress on the second syllable of class 2 words: when a three-syllable class 2 word, such as *màtànál* ‘gift’, occurs at the end of the IP (either in isolation or at the end of the sentence), it receives a boundary tone (T%) on its final syllable that cancels the effect of OCP(H) on the previous syllable, as argued in §5.5. Crucially, in that context the second syllable surfaces as accented: [mà.**tá**.nal$^{T\%}$], as in (21), confirming its foot head status, despite its contiguity with the following foot head.

<table>
<tr><td></td><td colspan="2">Underlying</td><td></td><td colspan="3">T%-insertion</td><td></td><td colspan="3">Parsing</td><td></td><td colspan="3">H-insertion</td></tr>
<tr><td>(21)</td><td>/ma</td><td>.ta.nal/</td><td>→</td><td>/ma</td><td>.ta</td><td>.nal/</td><td>→</td><td>(ma</td><td>.ˈta)</td><td>(ˈnal)</td><td>→</td><td>(ma</td><td>.ˈta)</td><td>(ˈnal)</td></tr>
<tr><td></td><td>|</td><td></td><td></td><td>|</td><td></td><td>|</td><td></td><td>|</td><td></td><td>|</td><td></td><td>|</td><td>|</td><td>|</td></tr>
<tr><td></td><td>L</td><td></td><td></td><td>L</td><td></td><td>T%</td><td></td><td>L</td><td></td><td>T%</td><td></td><td>L</td><td>H</td><td>T%</td></tr>
</table>

Besides initial L tones, the inherently accented syllable of some suffixes (cf. §3.3) constitutes the other instance of tone-driven stress in our proposal. We explain that pre-defined accent as a lexical H tone linked to that syllable. For instance *jpàstútik* ‘we (exclusive) make it’, featuring the accented suffix -*tútik* ‘1st person plural exclusive’ (cf. §3.3, (7b)), is analyzed as in (22): the H tone on the second syllable forces a parsing that makes it coincide with a foot head (tone-driven parsing). The remaining first syllable is parsed as a monosyllabic foot, which surfaces without a H tone because of OCP(H).

	Underlying				Parsing				H-insertion				OCP(H)		
(22)	/jpas	.tu	.tik/	→	(ˈjpas)	(ˈtu	.tik)	→	(ˈjpas)	(ˈtu	.tik)	→	(ˈjpas)	(ˈtu	.tik)
		|				|			|	|			ǂ	|	
		H				H			H	H			H	H	

It is necessary to assume that the initial syllable /jpas/ in (22) gets footed and as such is assigned a H tone in the course of the derivation, because an immediately preceding H-toned syllable gets automatically deaccented (i.e. loses its H tone). This happens for instance when *jpàstútìk* is preceded by the (accented) negation *mú*: *mú* + *jpàstútìk* → ***mù*** *jpàstútìk* ‘we (excl.) don’t make it’ (cf. 3.2, example (6)). If /jpas/ remained unfooted, it would not deaccent a preceding accented syllable.

Interpreting deaccenting as the effect of OCP(H) was noted in §5.5 as one of the interests of a tonal analysis on monosyllables. To posit H tones on polysyllables allows to extend this explanation to the deaccenting of accented syllables such as in (22), but also at the end of any polysyllabic word. For instance, the final syllable of *pòkó* ‘old’ gets deaccented when preceding an accented syllable, such as *ná* ‘house’ in (23a), analyzed as in (23b).

(23) a. pòkó + ná → pò**kò** ná ‘old house’

	H-insertion				OCP(H)		
b.	(po	.ˈko)	(ˈna)	→	(po	.ˈko)	(ˈna)
	|	|	|		|	ǂ	|
	L	H	H		L	H	H

Finally, there are reasons to think that no default L-insertion takes place on foot non-heads parallel to the H-insertion on foot heads. This can be gathered from the behavior of -σσ́ accented suffix such as -*etík* ‘plural’. Their first syllable is normally unaccented, e.g. in *mís-**è**tík* ‘cats’ or *pókòk-**è**tík* ‘toads’. We analyze this last word as in (24): the inherent H tone on the fourth syllable forces a parsing whereby it begins a foot, which leaves the third syllable as a degenerate foot, whose H tone is canceled by the OCP(H).

Underlying Parsing H-insertion OCP(H)

(24) /po.ko.ke .tik/ → (ˈpo.ko)(ˈke) (ˈtik) → (ˈpo .ko) (ˈke) (ˈtik) → (ˈpo .ko) (ˈke) (ˈtik)

| | | | | | ǂ |

H H H H H H H H

Interestingly, when such a word occurs at the end of the IP and the final syllable consequently conveys the phrasal pitch excursion, then the penult gets accented: [pó.kò.**ké**.tik$^{T\%}$]. The penultimate accent here confirms that the third syllable (ke) is a foot head, supporting the parsing proposed in (24). In particular, granted the possibility of iambic inversion invoked above (cf. (19)), one could have considered an alternative parsing with a final iamb: (ˈpo.ko)(ke.ˈtik). This parsing is ruled out, as it implies no accent should arise on the third syllable at the end of the IP, contrary to the facts. In our account, a boundary tone (T%) replaces the inherent H tone of the suffix, and so the previous syllable keeps its metrical H tone, as it is no longer blocked by the OCP(H), as shown in (25). This confirms that that syllable is not L-toned, as it contrasts with the initial syllable of class 2 disyllables such as *pòkó* 'old', which remains unaccented when the word occurs at the end of the IP: [**pò**.ko$^{T\%}$] (cf. §6.2, Table 10).

Underlying T%-insertion Parsing H-insertion

(25) /po.ko.ke .tik/ → /po.ko.ke .tik/ → (ˈpo.ko)(ˈke .tik) → (ˈpo .ko) (ˈke .tik)

| | | | | |

H H T% T% H H T%

This phenomenon proves that that particular syllable is underlyingly unspecified for tone. We conjecture that all non-initial unaccented syllables in general are similarly underlyingly unspecified for tone and phonetically interpreted as L-toned.

To summarize, the analysis we've sketched here develops the idea that CT accent patterns have a fundamental tonal component, combined with metrical stress. The relationship between both is mediated on the one hand by the mutual repulsion between (initial) L tones and foot heads, which forces a marked iambic footing when another syllable is available (i.e. in polysyllabic words); on the other hand, it is also mediated by the mutual attraction between H tones and foot heads, which forces a parsing where the H-toned syllable starts a trochee. Additionally, toneless foot heads receive a metrical H tone, which must comply with the OCP(H), whereas toneless non-heads remain toneless and are phonetically interpreted as L-toned. In that sense, our analysis does not entail an exhaustive tonal specification; in fact, it places CT on the low range of the tonal density scale (Gussenhoven 2004). The fact that boundary tones override both lexical and metrical tones on IP last syllables has been particularly useful to confirm some aspects of our analysis. We conclude with a synthesis of the relationship between our initial descriptive accented/unaccented distinction and our final categories in terms of tone and stress in Table 14.

TABLE 14
ACCENT VS STRESS+TONE

Initial description	Analysis	
Accent	Stress	Tone
Accented	+	H (underlying or metrical)
Unaccented	+	H cancelled by OCP(H)
	+	L
	-	L
	-	toneless

8. Discussion

The tonal/metrical analysis we just sketched has both strengths and weaknesses. On the positive side, it deals well with the prosody of monosyllables, provides a satisfactory account for polysyllables and offers an elegant explanation for the interaction between phrasal and word-level prosody. On the other hand, it involves a high degree of redundancy between stress and metrical tone, as accented syllables are explained both as foot heads and as H-toned.

To be sure, a pure stress account is a possible alternative, where accented and unaccented would simply correspond to stressed and unstressed, without the need of tone. Such an account would readily explain the prosody of class 1 words. Likewise, deaccenting would be simply interpreted as (anticipatory) stress deletion in case of stress clash.

As for class 2 words, their unaccented/unstressed first syllable or foot could be analyzed as a case of extrametricality, that is, the fact that some segment or prosodic constituent (syllable, foot) behaves as invisible for the rules creating metrical structures in a language (Liberman and Prince 1977, Hayes 1982, 1995, Halle and Vergnaud 1987). Extrametricality is especially appealing when stress is calculated from one of the word's edges but appears deeper inside the word than a simple rule would explain. Making some elements at the edge extrametrical then allows the rule to simply skip them (e.g. antepenultimate stress in Latin, Macedonian and many other languages is dealt with by marking the last syllable as extrametrical and constructing a final trochee, Hayes 1995; initial extrametricality of both syllables and feet has also been proposed in Kashaya: Buckley 1994). Returning to CT class 2 words, initial extrametricality, through a diacritic on the initial syllable, could explain the unstressed first syllable of disyllables and the unstressed first foot of longer words.

The challenge for a pure stress account lies in class 2 monosyllables: under this approach, the full stress domain (the only syllable) would end up extrametrical, as they remain unaccented. That is, those would be unstressable words. However, most authors agree that this should be strongly banned: according to the so-called *non-exhaustiveness condition* (Franks 1989), extrametricality cannot exhaust the domain to which it applies (Hayes, 1982, 1991; Prince, 1983; Inkelas, 1989; *inter alias*). For one thing, languages featuring edge extrametricality in their stress system anyway display stress on monosyllables. If in CT unaccented means unstressed, then class 2 monosyllables falsify the non-exhaustiveness condition, which is problematic. Additionally, if they were unstressed, they should be consistently shorter than their accented relatives, considering that duration appears as a correlate of stress in polysyllables. But, as shown in Table 7 (see §5.4 above), duration does not correlate systematically with the accented/unaccented distinction in monosyllables. This justifies postulating a hybrid stress/tone system such as ours instead of a pure stress approach: under our proposal, class 2 monosyllables are stressed but L-toned.

9. Conclusion

This is the first in-depth study of CT word-level prosody. On the descriptive side, we have established the existence of two lexical accent classes and shown how accented and unaccented syllables are distributed over morphemes, how they interact between themselves (deaccenting) and with phrasal intonation (final neutralization). On the analytical side, we have sketched a possible explanation based on a H/L tonal opposition and stress, namely a syllabic trochaic metrical pattern, with exhaustive parsing and allowing monosyllabic feet, and tentatively a main stress on the leftmost accented (i.e., H-toned) syllable. Our proposal is minimally tonal at the lexical level: L tones can only appear on the initial syllable of roots, whereas lexical H tones only appear on some long suffixes; other syllables are underlyingly toneless. Additionally, all foot heads potentially receive a metrical H tone. The cornerstone of our proposal is the existence of unaccented (i.e., non-H-toned) monosyllables, with contrastive low pitch, which make a tonal approach preferable to a pure stress one. In sum, CT constitutes an interesting new case of a language with interacting stress and phonological tone, feeding the typology of prosodic systems (Hyman 2006, 2009; Jun 2014).

References

AISSEN, JUDITH. 1987. Tzotzil clause structure. Dordrecht: Reidel.

AISSEN, JUDITH. 2017. Special clitics and the right periphery in Tsotsil. On looking into words (and beyond): Structures, Relations, Analyses, eds. Claire Bowern et al., Berlin: 235–262. Language Science Press.

ARELLANES ARELLANES, FRANCISCO. 2009. El sistema fonológico y las propiedades fonéticas del zapoteco de San Pablo Güilá. Descripción y análisis formal. Ph.D. dissertation, Colegio de México.

AVELINO, HERIBERTO; EURIE SHIN; AND SAM TILSEN. 2011. The phonetics of laryngealization in Yucatec Maya. New Perspectives in Mayan Linguistics, ed. Heriberto Avelino, 1–20. Newcastle: Cambridge Scholars Publishing.

BENNETT, RYAN. 2016. Mayan phonology. Language and Linguistic Compass 10:469–514.

---. 2018. Recursive prosodic words in Kaqchikel (Mayan). Glossa 3(1):67. 1–33.

BENNETT, RYAN AND ROBERT HENDERSON. 2013. Accent in Uspanteko. Natural Language and Linguistic Theory 31(3):589–645.

BENNETT, RYAN, ROBERT HENDERSON, AND MEGAN HARVEY. 2019. The interaction of tone and intonation in Uspanteko. Proceedings of the 19th international congress of phonetic sciences, Melbourne, 452–456.

BENNET, RYAN, ROBERT HENDERSON, AND MEG HARVEY. 2022a. Tonal variability and marginal contrast: Lexical pitch accent in Uspanteko. Prosody and prosodic interfaces, eds. Haruo Kubozono, Junko Ito, and Armin Mester. Oxford, UK: Oxford University Press.

BENNET, RYAN, MEG HARVEY, ROBERT HENDERSON AND TOMÁS ALBERTO MÉNDEZ LÓPEZ. 2022b. The phonetics and phonology of Uspanteko (Mayan). Language & Linguistics Compass 16 (9), e12467. https://doi.org/10.1111/lnc3.12467.

BERLIN, BRENT. 1962. Esbozo de la Fonología del Tzeltal de Tenejapa, Chiapas, Estudios de Cultura Maya, 2, pp. 17–36.

BICKMORE, LEE. 1995. Tone and stress in Lamba. Phonology 12: 307–341.

BOERSMA, PAUL & DAVID WEENINK. 2020. Praat: Doing phonetics by computer [Computer program]. Version 6.1.16. <http://www.praat.org/>

BROWN, PENELOPE. 1996. Isolating the CVC Root in Tzeltal Mayan: A Study of Children's First Verbs. Proceedings of the 28th Annual Child Language Research Forum, ed. Eve V. Clark, 41–52. Stanford, CA: CSLI/University of Chicago Press.

BUCKLEY, EUGEN. 1994. Persistent and cumulative extrametricality in Kashaya. Natural Language and Linguistic Theory 12: 423–464.

CAMPBELL, LYLE. 2017. Mayan history and comparison. The Mayan Languages, eds. Judith Aissen, Nora England and Roberto Zavala Maldonado, 43–61. London/New York: Routledge.

COHEN, ANTONIE, AND LOHAN 'T HART. 1967. On the anatomy of intonation. Lingua 19: 177–92.

COWAN, MARION M. 1969. Tzotzil grammar. Mexico City: Summer Institute of Linguistics.

CUTLER, ANNE. 2005. Lexical stress. The handbook of speech perception, eds. David Pisoni and Robert Remez, 264–289. Malden: Blackwell.

DE LACY, PAUL. 2002. The interaction of tone and stress in Optimality Theory. Phonology 19:1–32.

DICANIO, CHRISTIAN, JOSHUA BENN AND REY CASTILLO GARCÍA. 2018. The phonetics of information structure in Yoloxóchitl Mixtec. Journal of Phonetics 68: 50–68.

ENGLAND, NORA; AND BRANDON BAIRD. 2017. Phonology and phonetics. The Mayan Languages, eds. Judith Aissen, Nora England and Roberto Zavala Maldonado, 175–200. London/New York: Routledge.

FAYTAK, MATTHEW, AND ALAN C. L. YU. 2011. A typological study of the interaction between level tones and duration. Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong: 659–663.

FRANKS, STEVEN. 1989. The Monosyllabic Head Effect. Natural Language and Linguistic Theory 7: 551–563.

FRY, DENNIS. 1955. Duration and intensity as physical correlates of linguistic stress. The Journal of the Acoustical Society of America 27 (4): 765–768.

GORDON, MATTHEW. 2002. A factorial typology of quantity-insensitive stress. Natural Language & Linguistic Theory 20: 491–552.

---. 2011. Stress: Phonotactic and Phonetic Evidence. The Blackwell Companion to Phonology, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice, 924–948. Malden, Mass.: Wiley-Blackwell.

---. 2014. Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. Word Stress: Theoretical and Typological Issues, ed. Harry van der Hulst, 83–118. New York: Cambridge University Press.

GUSSENHOVEN, CARLOS. 1983. Focus, mode, and the nucleus. Journal of Linguistics, 19(2): 377–417.

---. 2004. The Phonology of Tone and Intonation. Cambridge: Cambridge University Press.

HALL, KATHLEEN CURRIE, ELIZABETH HUME, T. FLORIAN JAEGER AND ANDREW WEDEL. 2018. The role of predictability in shaping phonological patterns. Linguistics Vanguard 4 (s2).

HALLE, MORRIS AND JEAN-ROGER VERGNAUD. 1987. An Essay on Stress. Cambridge, Massachusetts: MIT Press.

HART, JOHAN ’T. 1981. Differential sensitivity to pitch distance, particularly in speech. Journal of the Acoustical Society of America 69(3): 811–821.

HAVILAND, JOHN B. 1981. Sk’op Sotz’leb: el Tzotzil de San Lorenzo Zinacantan. Mexico: UNAM.

HAYES, BRUCE. 1995. Metrical Stress Theory, Principles and Case Studies. Chicago: The University of Chicago Press.

HERNÁNDEZ GREEN, NESTOR. 2015. Morfosintaxis verbal del otomí de acazulco. Ph.D. dissertation, Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS).

HERRERA ZENDEJAS, ESTHER. 2013. Patrón acentual, F0 y consonantes laríngeas en el tsotsil de Venustiano Carranza. Paper presented at CILLA VI, October 24-26, University of Texas at Austin.

---. 2014. Patrones fónicos del tsotsil. Mapa fónico de las lenguas mexicanas: formas sonoras 1 y 2, ed. Esther Herrera, 337–383. Mexico: El Colegio de México.

HUALDE, JOSÉ IGNACIO, OIHANA LUJANBIO AND FRANCISCO TORREIRA. 2008. Lexical tone and stress in goizueta basque. Journal of the International Phonetic Association 38 (1): 1–24.

HULST, HARRY VAN DER. 2011. Pitch Accent Systems. The Blackwell Companion to Phonology, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice, 1003–1026. Malden, Mass.: Wiley-Blackwell.

---. 2014. The study of word accent and stress: past, present, and future. Word Stress: Theoretical and Typological Issues, ed. Harry van der Hulst, 3–55. New York: Cambridge University Press.

HYMAN, LARRY. 2006. Word-prosodic typology. Phonology 23, 225–257.

---. 2009. How (not) to do phonological typology: the case of pitch-accent. Language Sciences 31:213–238.

HYMAN, LARRY AND KEMMONYE C. MONAKA. 2011. Tonal and Non-Tonal Intonation in Shekgalagari. Prosodic categories: Production, Perception and Comprehension, eds. Sónia Frota, Gorka Elordieta and Pilar Prieto, 267–289. Springer.

JUN, SUN-AH. 2014. Prosodic Typology. Prosodic Typology: The Phonology of Intonation and Phrasing, ed. Sun-Ah Jun, 430–458. Oxford: Oxford University Press.

KAUFMAN, TERRENCE. 1971. Tzeltal Phonology and Morphology. Berkeley: University of California Press.

---. 1972. El proto-tzeltal-tzotzil: fonología comparada y diccionario reconstruido. Mexico: Centro de Estudios Mayas, Cuaderno 5, UNAM.

LAUGHLIN, ROBERT M. 1975. The great Tzotzil dictionary of San Lorenzo Zinacantán. Washington: Smithsonian Institution Press.

LIAN-HEE, WEE. 2019. Phonological tone. Cambridge: Cambridge University Press.

LIBERMAN, MARK, AND ALAN PRINCE 1977. On Stress and Linguistic Rhythm. Linguistic Inquiry 8 (2): 249–336.

LÓPEZ MENDOZA, EDUARDO DE JESÚS. 2022. La definitud y las funciones del enclítico =e en el tsotsil sureño. MA dissertation, CIESAS.

MADDIESON, IAN. 1978. Universals of tone. Universals of Human Language, volume 2: Phonology, ed. Joseph Greenberg et al., 335-363. Stanford: Stanford University Press.

MARTÍNEZ-PARICIO, VIOLETA AND RENÉ KAGER. 2016. Metrically conditioned pitch and layered feet in Chugach Alutiiq. Loquens 3(2), e030. doi: http://oadoi.org/10.3989/loquens.2016.030.

MATEO TOLEDO, ELADIO. 2008. The family of complex predicates in Q'anjob'al (Maya); their syntax and meaning. Ph.D. dissertation, University of Texas at Austin.

MEREDITH, SCOTT. 1990. Issues in the phonology of prominence. Ph.D. dissertation, MIT.

ODDEN, DAVID. 1986. On the role of the Obligatory Contour Principle in phonological theory. Language 62: 353–383.
PIERREHUMBERT, JANET, AND MARY BECKMAN. 1988. Japanese tone structure. Cambridge: MIT Press.
PIKE, EUNICE V. 1951. Tonemic-intonemic correlation in Mazahua (Otomi). International Journal of American Linguistics 17:37-41.
PITTAYAPORN, PITTAYAWAT. 2018. Phonetic and systemic biases in tonal contour changes in Bangkok Thai. Tonal change and neutralization, eds. Haruo Kubozono and Mikio Giriko, 249–278. Berlin: Mouton de Gruyter.
POLIAN, GILLES. 2013. Gramática del tseltal de Oxchuc. Mexico: Centro de Investigaciones y Estudios Superiores en Antropología Social.
POLIAN, GILLES. 2017. Tseltal and Tsotsil. The Mayan Languages, eds. Judith Aissen, Nora England and Roberto Zavala Maldonado, 610–647. London/New York: Routledge.
SARLES, HARVEY. 1966. A Descriptive Grammar of the Tzotzil Language as Spoken in San Bartolomé de los Llanos, Chiapas, México. Ph.D. dissertation, University of Chicago.
SHKLOVSKY, KIRILL. 2011. Petalcingo Tseltal intonational prosody. Proceedings of Formal Approaches to Mayan Linguistics (FAMLi), eds. Kirill Shklovsky, Pedro Mateo Pedro and Jessica Coon, 209–220, Cambridge, MA: MIT Working Papers in Linguistics.
STEVENS, KENNETH N. 2000. Acoustic Phonetics. MIT Press.
WHALEN, D. H. AND YI XU. 1992. lnformation for Mandarin Tones in the Amplitude Contour and in Brief Segments. Phonetica 49: 25–47.
YIP, MOIRA. 2002. Tone. Cambridge: Cambridge University Press.
ZEE, ERIC. 1978. Duration and intensity as correlates of F0. Journal of Phonetics 6: 213–220.
ZEC, DRAGA. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. Phonology 16: 225–264.