Acoustic Correlates of Prominence in Kala Lizu (Tibeto-Burman)

Katia Chirkova
CNRS-CRLAO, Nankai University
Paris, Tianjin

Pei-Yu Hou
North Carolina State University
Raleigh

Rada Chirkova
North Carolina State University
Raleigh

Angélique Amelot
CNRS-LPP
Paris

Abstract

This study investigates acoustic correlates of prominence in Lizu (Tibeto-Burman). Lizu has been argued to have a hybrid prosodic system combining lexical tone on monosyllabic words and prominence patterns with stress-like and tonal characteristics on polysyllabic words, although empirical evidence is lacking. This study presents an acoustic investigation of the pitch patterns on disyllabic words in the Kala variety of Lizu (HL, HH, LH). Using the parameters of duration, intensity, and f0, it attempts to sort out different forms of prominence, and to explore their interaction with lexical tone. The measurements are taken from experimental data with eight Lizu speakers (4 male and 4 female). The acoustic results and statistical analyses suggest that the first syllable in Kala Lizu is the position of prominence in all three pitch patterns. Our results are consistent with interpreting the pitch pattern HL as more stress-like (with intensity as an acoustic correlate of stress) and the pitch patterns HH and LH as more tone-like (with stress cued by full realization of lexical tone). This study contributes to a better understanding of the prosodic organization of Kala Lizu and it also suggests methodology for further exploration of other Lizu varieties.

Keywords

Prosodic prominence   Stress   Tone   Lizu   Tibeto-Burman

1. Introduction

This study examines the prosodic organization of Lizu (ISO-639 code *ers*), an underdescribed Tibeto-Burman (Southern Qiangic) language. Lizu is a continuum of closely related linguistic varieties which are spoken in three counties of Sichuan Province in the People's Republic of China: Jiulong (九龙县, Written Tibetan [WT] *brgyad zur*), Muli Tibetan Autonomous County (木里藏族自治县, WT *smi li rang skyong rdzong*, hereafter Muli), and Mianning (冕宁县). Lizu is a tonal language, in which morphemes are generally monosyllabic, but words are mostly disyllabic and composite. Pioneering descriptions of various Lizu varieties (Huang & Renzeng 1991; Ikeda 2009; Yu 2012; Chirkova & Chen 2013) suggest that they are very similar and differ mainly in terms of word-level prosody. All reported varieties have simple tone systems, consisting of a high-low contrast (H vs. L) on monosyllabic words and a limited number of possible surface pitch patterns over polysyllabic words (see Yu 2012, 6–11 for discussion). These patterns represent various combinations of high and low tones (e.g., HH, HL, LH, HHH, HHL, LHL, etc.) and show

Acknowledgements We would like to thank our Lizu consultants Wang Xuecai 王学才, Wang Tshi'er Zhuoma 王次尔卓玛, Song Changxiu 宋长秀, Wang Sifeng 王四风, Wang Zi'an 王字安, Song Chaoqun 宋超群, Huang Xiuying 黄秀英, Linqing Lamu 林青拉姆, and Huang Wujin 黄五斤 for their patience, enthusiasm, and willingness to facilitate this study. Earlier versions of this paper were presented at research seminars at the School of Literature, Nankai University (Tianjin, China) on December 12, 2018, and at Helsinki University (Helsinki, Finland) on December 19, 2019. We are thankful for the feedback we received following those presentations. We are grateful to the Journal editors and the anonymous reviewers for their detailed reviews and valuable suggestions, which helped to improve the manuscript. We would also like to thank Professor Yeshes Vodgsal Atshogs for translating the abstract into Chinese. We gratefully acknowledge the financial support of the National Science Foundation (grant number 1747555) to Rada Chirkova. This work is partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program "Investissements d'Avenir" (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris-ANR-18-IDEX-0001.

complex correspondences between varieties. This is illustrated in Table 1 with data from two varieties of Lizu, spoken in Kala 卡拉 Township, Muli County (Muli Lizu 1 & Muli Lizu 2); a variety of He'ai 和爱 Township in Mianning county (Mianning Lizu); and a variety of Naiqu 乃渠 Township in Jiulong county (Jiulong Lizu).[1] Examples in Table 1 are limited to disyllabic words because they represent the most common type of polysyllables.

Table 1 Surface pitch patterns on disyllabic words uttered in isolation in four varieties of Lizu

Gloss	Muli Lizu 1	Muli Lizu 2	Mianning Lizu	Jiulong Lizu
head	wuli HL	wuli LH	vʉli HH [HH~HL][2]	wuli LH
eye(ball)	Ndosə HL	nduose LH	ʒdosɨ LL [LH]	dosɯ LH
nail	ledzə HL	letsa LH	lidzɑ 'claw' HH [HH~HL]	letsa LH
thin	bibje HH	bibi HH	bibi HH [HH~HL]	bibi LH
year	dʒutʃʰe HH	dʐutʂʰɯ HL	dʑitʂʰɤ LL [LH]	titɕʰə HH
rabbit	midzə HH	mitsə LH	ɲidzɨ HH [HH~HL]	
ear	nɐpi LH	napi HH	nɑpi HH [HH~HL]	napu LH
tree	sepu LH	sepu LH	sɨpʉ LL [LH]	səpu LH
long	ʃɐʃɐ LH	ʂaʂa HH	ʂiʂɑ LL [LH]	ʂaʂa LH

The sample in Table 1, albeit limited, illustrates the three main differences in surface pitch patterns on disyllabic words in various varieties of Lizu. First, different number of pitch patterns per variety, e.g., three in Muli Lizu (HL, HH, LH) versus two in Mianning and Jiulong Lizu (HH, LH). Second, different ratios of pitch patterns per variety. For example, the three pitch patterns in Muli Lizu 1 are relatively evenly distributed across the lexicon (see Section 2.1), whereas the majority of disyllabic words in Jiulong Lizu (as provided in Ikeda 2009) appear to have the pitch pattern LH. Third, complex correspondences between pitch patterns across varieties. The three surface pitch patterns in disyllabic words in Muli Lizu 1 correspond rather irregularly to the three surface pitch patterns in Muli Lizu 2, and the two surface pitch patterns in Mianning Lizu and Jiulong Lizu.

In their work on the ʃætɕʰopæ HHH variety of Lizu (Muli Lizu 1 in Table 1) (hereafter Kala Lizu), Chirkova & Chen (2013) observe that syllables in disyllabic words differ in prominence, manifested with pitch and durational cues. In their analysis, the three surface pitch patterns HL, HH, and LH in that variety are described, respectively, as "left-prominent", "equally-prominent", and "right-prominent". They further suggest that Kala Lizu may have a hybrid prosodic system that combines lexical tone on monosyllabic words and prominence patterns with stress-like and tonal characteristics on polysyllabic words. This proposal is of interest in the context of differences in the pitch patterns on disyllabic words among Lizu varieties, because it may provide a possible explanation for these differences. Namely, recognition of stress in addition to tone opens up the possibility that the complex correspondences among Lizu varieties are due to differences in the structure and realization of prominence relations, as common among dialects of the same language and among related languages (see Smith, Erickson, and Savariaux 2019; Smith and Rathcke 2020 for recent accounts). However, such an explanation cannot be directly applied, for the prosodic organization of different Lizu varieties remains little studied, and even in relation to Kala Lizu, for which a hybrid prosodic system has been proposed, the proposal is based on auditory impressions and has not been experimentally tested. Furthermore, given that supporting evidence for the proposal in Chirkova & Chen (2013) was limited to words spoken in isolation, a context that potentially confounds word- and phrase-level phenomena (Gordon 2014; Roettger & Gordon 2017), surface pitch patterns on polysyllabic words in Kala Lizu likely represent complex interactions between lexical tone, possibly, word-level stress, and post-lexical tonal events associated with prominent syllables (pitch accents or phrase-level stress) and prosodic boundaries (boundary tones).

The goal of this study is to empirically examine the previous claims of stress in Kala Lizu. To that end, we explore a small corpus of disyllabic words in two contexts: citation (where the word is both phrase-final and bears the main stress of the phrase) and phrasal (where the word is both preceded and followed by other words so that it is shielded from edges of the phrase). We consider three cross-linguistically common acoustic correlates of prominence, that is, duration, (overall) intensity, and fundamental frequency (f0). On that basis, we attempt to sort out possible forms of prominence at a word level (word-level stress) and at a phrase level (phrase-level stress and boundary tones), and to assess their interaction with lexical tone. Our acoustic investigation reveals differences between the two contexts examined and among the three pitch patterns in each context. The observed differences between the contexts are likely due to the addition in citation context of post-lexical tonal events associated with prominent syllables (phrase-level stress) and prosodic boundaries. The main differences among the three pitch patterns include an opposition between the pitch pattern HL, on the one hand, and the pitch patterns HH and LH, on the other hand. For both types of patterns, we suggest that the first syllable is the position of prominence, albeit manifested through different acoustic parameters. These include intensity in the case of HL and full lexical tone realization in the case of HH and LH. It can therefore be suggested that Lizu has fixed stress on the first syllable in disyllabic words, and two types of words: with more stress-like characteristics (words with the surface pitch pattern HL) and with more tone-like characteristics (words with the surface pitch pattern HH and LH).

Our study contributes to a better understanding of the prosodic organization of Lizu and thereby to a more precise description of that language. It also has implications for the study of prosody of other Lizu varieties and suggests methodology for their further exploration. This study also contributes novel data and insights into hybrid prosodic systems that combine lexical tone and prosodic prominence.

The structure of this paper is as follows. The next section serves as the general background. It details some of the linguistic particulars of Kala Lizu and its prosody, including an overview of word stress claims in that variety (Section 2.1). Given the understudied nature of Lizu, a priori it is difficult to predict which correlates could be particularly indicative of word- and phrase-level in that language. For that reason, the next section also provides an overview of acoustic correlates of prominence in a cross-linguistic context (Section 2.2). Section 2 is concluded by this study's research questions and predictions formulated on the basis of sections 2.1 and 2.2 (Section 2.3). Section 3 describes the experimental methods of the study reported on in this paper. Section 4 presents the results, which are further discussed in Section 5. A conclusion (Section 6) rounds off the paper.

2. Tone and prominence in Kala Lizu and in a cross-linguistic context

2.1. Linguistic particulars of Kala Lizu and its prosody

Kala Lizu is an isolating (analytic) language with no inflectional morphology and little productive derivational morphology. It is phonologically and morphologically monosyllabic and tonal, with the majority of words being disyllabic and composite.[3] Compounding, affixation, and reduplication are the three main word formation processes. Kala Lizu is mostly suffixing on nouns and prefixing on verbs. Verbal prefixes are synchronically productive, whereas nominal affixes are not. Reduplication is part productive and part lexically idiosyncratic. In disyllabic words, full reduplication is found (that is, repetition of an entire root without phoneme changes or additions).

The syllable structure is (N)(C1)(C2)V^T, where N is nasal (homorganic to the following consonant); C1 can be any initial consonant (that is, *p pʰ b t tʰ d k kʰ g q qʰ ɢ ts tsʰ dz tʃ tʃʰ dʒ tɕ tɕʰ dʑ m n ɲ ŋ s z ʃ ʒ ɕ ʑ x ɣ ɦ̃ w ɹ j l ɬ*); C2 can be one of the following set: *j w ɹ z ʑ ɕ ʃ*; V can be any vowel (that is, *i y u e ə o æ ɐ*); T is tone; and parentheses indicate optional consonants (see Table 2 for examples). Vowels do not contrast in length. Most syllables have a simple structure C1V^T.

On monosyllabic words, Kala Lizu distinguishes two lexical tones, which in citation form, are realized as low rising, followed by a short fall; and high falling. This is a robust opposition, supported by many minimal pairs. Tones are fully independent, that is to say that any vowel may bear any tone and all consonants, including sonorants and voiced obstruents, can be associated with both tones. Table 2 provides some illustrative examples, organized by vowel frontness and height.

Table 2 Minimal pairs for the two lexical tones on monosyllabic words in Kala Lizu

Low Rising	High Falling
ne ‘two’	ne ‘you, thou’
tæ ‘mule’	tæ ‘tendon’
tʃə ‘frost’	tʃə ‘to bite’
ŋu ‘to weep’	ŋu ‘cow’
mo ‘again’	mo ‘tomb’
Ndzɐ ‘drum’	Ndzɐ ‘Chinese’

The mean pitch tracks for the two contrastive tones on monosyllabic words are plotted in Figure 1 on the basis of 26 tokens (words in citation form, with a mix of onsets in terms of voicing and sonority), with 2 repetitions for each lexical item.[4]

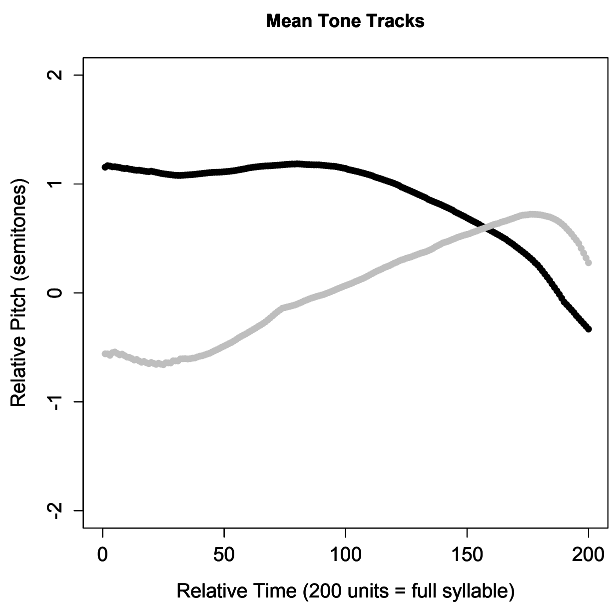


Figure 1 Mean tone tracks for the two contrastive tones on monosyllabic words, as uttered in isolation: (1) low rising (in gray), 11 tokens; (2) high falling (in black), 15 tokens. Plotted as relative time versus Hz. Normalized for duration and mean low rising pitch

In running speech, the two tones are realized as level, and even in citation form, realizations without a perceptually salient f0 fall or rise are also acceptable for native speakers. For these reasons, Kala Lizu can be analyzed as having two basic tonal categories H and L, so that the surface contour tones on monosyllabic words in citation form can be decomposed into the lexical tones H and L, followed by a post-lexical (boundary) tone with a surface high falling contour (HL%).[5] The processes involved can be presented as follows: (1) H+HL% > [HL] for the high tone, and (2) L+HL% > [LHL] for the low tone.

In disyllabic words in citation form three word-level surface pitch patterns (hereafter pitch patterns) are observed. In the analysis of Chirkova & Chen (2013), these three pitch patterns are described as follows:

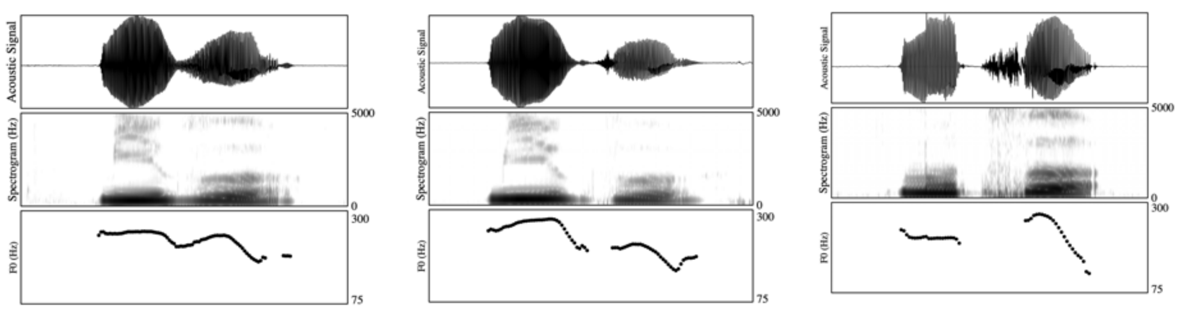
(1) Equally-prominent pattern: Both syllables have a high, slightly falling pitch contour, and the second syllable has a high falling pattern. Both syllables sound equally prominent and no special lengthening is observed on either syllable (despite the presence of final lengthening due to the fact that these words were elicited in isolation as an utterance on its own).

(2) Left-prominent pattern: The f0 peak is typically realized before the end of the first syllable, where the pitch starts to fall and continues to fall in the second syllable. The duration of the second syllable, despite final lengthening, is relatively shorter and the syllable also sounds less prominent.

(3) Right-prominent pattern: Within the first syllable, there is a low-level pitch contour with a slight rise. The f0 peak is realized within the second syllable where there is also a clear fall. The duration of the second syllable is relatively longer than that of the first syllable.

Chirkova & Chen (2013) also note that Kala Lizu does not exhibit vowel reduction in non-prominent positions. Based on the location of the f0 peak within the word, the three pitch patterns will hereafter be referred to as HH, HL, and LH, respectively.

A three-way contrast can be observed in Figure 2 over the words midzə HH 'rabbit', midzə HL 'pepper', and mutsə LH 'cat', which have a comparable segmental composition.



midzə HH 'rabbit' midzə HL 'pepper' mutsə LH 'cat'

Figure 2 Waveform, spectrogram and f0 contour for the three-way pitch pattern contrast over the disyllabic words midzə HH 'rabbit' (left), midzə HL 'pepper' (middle), and mutsə LH 'cat' (right) by a female speaker

The three pitch patterns on disyllabic words are distributed relatively evenly throughout the lexicon. In our corpus of basic lexical items containing 1,855 disyllabic words: 610 words (or 33%) have the pitch pattern HH, 610 words (or 33%) have the pitch pattern HL, and 635 words (or 34%) have the pitch pattern LH. At the same time, no minimal three-way contrast for the pitch patterns has been attested.[6] Attested contrasts are binary and they vary slightly depending on word-formation type. Over nouns, contrasts between the HH and LH patterns are slightly more common. Examples include: tsʰomo HH person-old 'old man' vs. tsʰomo LH person-grave 'tomb'. Over verbs, contrasts between the pitch patterns HL and LH are slightly more common. Examples include: ne-ko HL downward-wither 'to wither, shrivel' vs. ne-ko LH downward-put 'to put inside'.

These tendencies can be further correlated with the productive sandhi rules in Kala Lizu. The first rule applies when monosyllabic morphemes are combined into words or compounds. According to this sandhi rule, (i) the tones on non-initial syllables are neutralized, and (ii) the tone of the initial syllable spreads onto the non-initial syllables.[7] Table 3 provides some illustrative examples.

Table 3 Examples of compounds composed of two monosyllabic words

	Ndʒə L ‘skin’	mu H ‘animal hair’
ŋu H ‘cow’	ŋuNdʒə HH ‘cow hide’	ŋumu HH ‘cow hair’
ɹwæ L ‘chicken’	ɹwæNdʒə LH ‘chicken skin’	ɹwæmu LH ‘chicken feathers’

Application of this sandhi rule provides a possible explanation for a somewhat more common occurrence of the contrast between the pitch patterns HH and LH over disyllabic nouns.

The second sandhi rule applies to verbs formed through prefixation.[8] If the verb root carries the low tone, the resulting tone of the prefixed verb is LH, as in dzə LH ‘eat’ > ne-dzə LH ‘eat up; have eaten’. If the verb root carries the high tone, the resulting pitch pattern of the prefixed verb is HL, as in dzə H ‘give birth’ > de-dzə HL ‘give birth; have given birth’. Application of this sandhi rule on verbs formed through prefixation provides a possible explanation for a somewhat more common occurrence of the contrast between the pitch patterns HL and LH over disyllabic verbs.

While the pitch patterns on some disyllabic words can be linked to the etymological tones of their constituent morphemes through the two sandhi rules above, in many cases, pitch patterns on disyllabic words are lexicalized and opaque (as in mjæ-mu HL eye-animal.hair ‘eyelash’, formed through compounding; based on the productive tone sandhi rule above, the expected pitch pattern for this word would be LH). All three pitch patterns are equally attested on words formed through compounding, affixation, and reduplication, as well as on monomorphemic disyllabic words. Put differently, none of the attested disyllabic word types in Kala Lizu is associated exclusively with only one pitch pattern or a subset of the three possible pitch patterns.[9] This is illustrated in Table 4, which provides contrastive sets for all possible types of pitch pattern contrasts (that is, HL vs. HH, HL vs. LH, and HH vs. LH), as attested in all possible disyllabic word types (that is, monomorphemic words and words formed through compounding, affixation, and reduplication).

Table 4 Pitch pattern contrasts on disyllabic words

Word formation type	**HL**	**HH**	**LH**
Monomorphemic	midzə ‘pepper’	midzə ‘rabbit’	
		kælæ ‘hardship’	kælæ ‘butterfly’
Compounding	mjæ-mu ‘eyelash’ (< mjæ L ‘eye’, mu L ‘animal hair’)	wo-mu ‘pig bristles’ (< wo H ‘pig’, mu L ‘animal hair’)	ɲo-dʒu ‘lambswool’ (< ɲo L ‘sheep’, dʒu L ‘wool’)
Affixation (with the female gender suffix -mæ)	ɲo-mæ ‘ewe’ (< ɲo L ‘sheep’)	wo-mæ ‘sow’ (< wo H ‘pig’)	ŋu-mæ ‘cow’ (< ŋu H ‘cow’)
(lexicalized)	ɲimæ ‘sun’		ɲimæ ‘living room’
Reduplication	tɐ-tɐ ‘to chew’	tɐ-tɐ ‘to hold in one’s arms’	
	be-be ‘to climb’		be-be ‘to crawl’

2.2. Acoustic correlates of prominence in a cross-linguistic context

In this study we explore the proposal that in addition to tone, that is, a featural property referring to contrastive relative pitch (e.g., Hyman 2001; Yip 2002), Kala Lizu has contrastive differences in the relative prominence of syllables in polysyllabic domains. The notion of prominence integrates discreteness, relative patterning, and hierarchical ordering; and it can be applied at different levels of linguistic analysis (see Cangemi & Baumann 2020 for a recent overview). These levels canonically include, on the one hand, the relative strength of syllables within a word (word-level stress), and, on the other hand, the relative strength of syllables within a breath group (or stress group, or phonological phrase) (phrase-level stress). Word-level stress is obligatory and culminative (Hyman 2006, 231; 2009, 216, 231), and it has a semantic or grammatical role. Phrase-level stress also has a culminative function; and it has a pragmatic role, relating to the focus and information structure of a sentence. Two more aspects of phrasal prominence that are of relevance in relation to the evidence in Chirkova & Chen (2013) (words in citation context, where the word constitutes an utterance of its own and is phrase-final) include: (i) the lengthening effects found at phrase edges, and (ii) boundary tones. Both have a delimitative function and, together with prominent syllables at a phrase level, are phonetic manifestations of the prosodic structure (e.g., Klatt 1975; Beckman & Pierrehumbert 1986; Hayes 1989; Byrd, Krivokapić & Lee 2006; Cho 2016).

The main acoustic correlate of tone is pitch (f0 as perceptual correlate). The acoustic correlates of prominence are more diverse and include duration, (overall) intensity, f0, spectral tilt, and vowel formant frequencies (especially F1 and F2) (e.g., Ladd 1996; Sluijter & van Heuven 1996; Sluijter et al. 1997; Crosswhite 2004; de Jong 2004; Hintz 2006; Fletcher 2010; Gordon 2011, 2014; Gordon & Roettger 2017). Of these, duration, intensity, and fundamental frequency have been recognized as acoustic correlates of prominence since pioneering work by Fry (1955, 1958), and have been extensively researched in various languages. The same set of physical properties (duration, intensity, f0) is also used for prominence effects found at phrase edges. Given this overlap in the use of acoustic correlates for prominence effects at the word- and the phrase- levels, the acoustic definition of word stress is heavily dependent on how these correlates are used on other levels of prosody in a given language.

A recent survey of 110 studies of 75 languages (Gordon & Roettger 2017) suggests duration as the most consistent correlate of word-level stress across languages. In most cases, the lengthening effect of stress is most consistently seen in the rhyme portion of the syllable (see Shen 1993; Sluijter & van Heuven 1995; de Jong and Zawaydeh 2002; Remijsen 2002; Chavez-Peon 2008). The effect of stress on intensity and f0, on the other hand, is less straightforward. At least in some languages, overall intensity and f0 are more correlated with phrase-level stress than with word-level stress (e.g., Bolinger 1958; Sluijter & van Heuven 1996; Campbell & Beckman 1997; Hintz 2006; see van Heuven 2018 for discussion).[10] Therefore, the combined effect of duration, intensity, and f0 may be a better way of capturing the perceptual prominence of stressed syllables than the latter two cues taken separately.

The acoustic definition of word-level stress also critically depends on the system of linguistic contrasts operating within a given language. For example, languages with phonemic length contrast in vowels tend to be less reliant on duration as acoustic correlate of stress. Finnish is a case in point, as its unstressed vowels may be phonetically longer than stressed vowels (see Suomi et al. 2001, 2003). Languages that combine lexical tone with word-level stress, such as Thai (Potisuk, Gandour & Harper 1996) and Pirahã (Everett 1998), tend to be less reliant on pitch for the realization of stress, because pitch is primarily used in the realization of lexical tone contrasts. Therefore, in tone languages with stress, prominence (at the word- and phrase- levels) is rather associated with segmental duration and intensity (e.g., Remijsen & van Heuven 2005; Chavez-Peon 2008; Hyman 2009; Guion et al. 2010; Caballero & Carroll 2015; Gordon & Roettger 2017).

In tone languages with word-level stress, tone and stress systems often interact. In many languages with sparse tonal specifications (that is, languages in which not every syllable has its own lexical tone), tonal distribution is dependent on word-level stress (see Pike & Oram 1976; Kristoffersen 2000; Gussenhoven 2004; Lahiri et al. 2005; Sun 2005, 2008; Chavez-Peon 2008; Lin 2009, 2012; Downing 2010; Caballero & Carroll 2015). Tonal specification is typically found only in stressed syllables, while unstressed syllables are unspecified for tone. In languages with dense tonal specification (that is, languages in which each syllable has its own lexical tone), interaction of tone and word-level stress may take more complex forms. Tonal languages canonically display full tone realization in stressed syllables; with tones also having wider f0 range. (The same strategy is also used at a phrase level to signal the difference between words in and out of focus, e.g., Shih 1988; Xu 1999; Chen & Gussenhoven 2008; Brunelle 2017). Unstressed syllables, on the other hand, may either maintain tonal contrasts (as in Thai, Potisuk, Gandour, & Harper 1996) or neutralize them (as in Burmese, Gruber 2011); and they may be additionally characterized by shortened duration, lesser intensity, and reduced vowel quality.

Finally, in some tonal languages with dense tonal specification, only some words can be analyzed as having stress. Beijing Mandarin is a case in point. It has a large number of words with neutral tone syllables, which are reduced in their segmental articulation (neutral tone, shortened duration, lesser intensity, centralized vowel). These words are commonly analyzed as having trochaic stress (Lin & Yan 1980). At the same time, the issue of whether disyllabic words, in which both syllables are specified for tone, also have stress is unsettled. Existing proposals include both trochaic and iambic stress patterns (e.g., Kratochvil 1964; Wang & Feng 2006).

2.3. Research goals, corpus, and predictions

Research goals: In this study, we experimentally explore the previous prominence claims in Lizu, as proposed in Chirkova & Chen (2013) in relation to words in isolation. We expect that surface pitch patterns on words in isolation (where the word is both phrase-final and bears the main stress of the phrase) represent complex interactions between lexical tone, possibly, word-level stress, and post-lexical tonal events associated with prominent syllables (phrase-level stress) and prosodic boundaries (boundary tones). To tease these factors apart, we compare the realization of words in isolation with the realization of the same words in frame sentences (in phrase-medial position, where the target word is both preceded and followed by other words so that it is shielded from edges of the phrase). Phrasal context is also expected to provide more straightforward evidence on word-level stress.

Corpus: One key consideration in constructing a corpus for this study was not to obscure the acoustic correlates of prominence duration and intensity by differences in the segmental make-up of the syllables. Specifically, the relatively complex syllable structure of Kala Lizu (with the possibility of syllable-initial clusters, (N)(C1)(C2), e.g., mb, bɹ, mbɹ, see Section 2.1) can potentially affect segment duration, because segments tend to be shorter in complex versus simple syllables (Fletcher 2010, 525). Vowel duration may also be influenced by vowel height (e.g. open vowels have inherently longer duration and greater intensity than close vowels) (e.g., Lehiste & Peterson 1959; Peterson & Lehiste 1960; Klatt 1975; van Santen 1992; Fletcher 2010).[11] For these reasons, we opted for words that would contain, on the one hand, same syllable onsets in both syllables (either simplex or cluster of a similar segmental composition, as in qɐtɐ HH 'knot' or NdzuNdzu HH 'button') and, on the other hand, same vowels in both syllables (see examples above), while relying as much as possible on attested minimal pairs in this language. Within these constraints, and given that different word classes do not affect tone patterns or tone sandhi rules (see Section 2.1), nominal and verbal reduplicated forms stood out as most suitable for our study, because they naturally present the advantage of the segmental control that facilitates obtaining acoustic evidence of asymmetrical prominence relations. The resulting corpus includes a total of 32 disyllabic lexical words (see Appendix). Of these, 25 are reduplicated verbs (including adjectives or intransitive stative verbs, e.g., be-be HL 'to climb', tɐ-tɐ HH 'to hold in one's arms', be-be LH 'to crawl'); and 7 are nous, of which 3 are reduplicated (NdzuNdzu HH 'button', tʃʰutʃʰu LH 'mouth harp', tʃʰɐtʃʰɐ LH 'magpie') and 4 are non-reduplicated, but have the same syllabic onsets and the same vowel in both syllables (qɐtɐ HH 'knot', kælæ HH 'harship', kælæ LH 'butterfly', jɐtɐ LH 'left').

Predictions: In our study, we examine three cross-linguistically common acoustic cues to prominence, that is, duration, intensity, and f0 (see section 2.2). Given that Kala Lizu has contrastive lexical tones on monosyllabic words, in line with the cross-linguistic tendencies discussed in Section 2.2, we assume that fundamental frequency, being the primary correlate of tone, is less readily available for conveying word-level stress distinctions, so that word-level prominence is more likely associated with duration and intensity.

If there is word-level stress in Kala Lizu, following again the cross-linguistic tendencies discussed in Section 2.2, it should be expected that (i) there are significant asymmetries in duration and intensity between the two syllables of the word; and (ii) there may also be asymmetry in the realization of tone contrasts between the two syllables of the word. In particular, tone contrasts in unstressed syllables may be reduced.

3. Methodology

This section describes the data collection procedure and the acoustic analysis carried out on Kala Lizu speech data.

3.1. Data collection procedure

We recorded the target words in two contexts: (i) citation and (ii) phrasal (phrase-medial). For the phrasal context, we used a number of frame sentences, (i) where the target words could be pronounced in the most natural way possible (see Appendix for details), and (ii) where the target word would not trigger special emphasis (contrastive focus). In sentence with verbs (including intransitive stative verbs), that was done by placing the target word in an utterance which contains a word that is explicitly focused (e.g., the adverb 'extremely'). It proved more difficult to ensure that the target word does not trigger special emphasis in sentences with nouns. Frame sentences for nouns included broad focus statements of the type 1SG ______ V (e.g., æ H ______ kʰe-Ndo=ɐ LH 'I saw ______'). However, given that the targets in such phrases represent new information, they can possibly attract post-lexical prominence (cf. Roettger & Gordon 2017). To minimize the potential influence of post-lexical prosody, sentences with nouns were randomized so that broad focus sentences that are minimally different are not listed next to each other.

Words on the word list were presented to the subjects to produce in response to an oral prompt (the meaning of the target word or sentence in Mandarin Chinese) in randomized order in the two contexts above. In each context, the target words were read each three to five times (with an average of four times).

Recordings were made in Qiaowa Town, the administrative seat of Muli county in December 2018. Speakers were recorded in a quiet room, using a Zoom H6 recorder (with a sampling rate of 16-bit at 48 kHz) and an AKG C520 headset microphone (which allowed to ensure a constant distance between the mouth of the speaker and the microphone so as to warrant the reliability of intensity as a correlate of stress, cf. Heuven 2018, 24; Kaland 2019, 58).[12]

The data were collected with the assistance of eight native speakers of Kala Lizu, of whom four were women (ages 40-65) and four were men (ages 32-69). Our language consultants were all fluent speakers of Lizu, using it in their daily lives, and who had spent most of their lives in Muli County. They had no reported history of speech or hearing problems. All language consultants recorded for this study received financial recompense for their time.

All data were recorded mono in uncompressed Waveform Audio File Format (.wav). The first author supervised the entire recording procedure. The duration of the collected recordings ranged between 15 and 22 minutes.

3.2. Data processing

The corpus was segmented manually in Praat (Boersma & Weenink 2020) with the assistance of a time-aligned spectrogram. For each target word, we marked the boundaries of its two vowels (as determined by the onset and offset of F2). An example is shown in Figure 3.

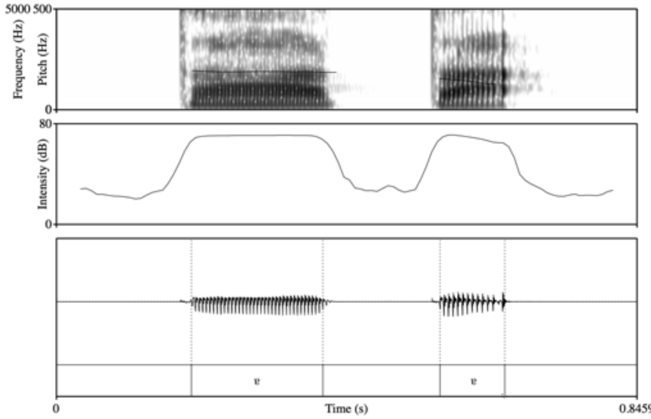


Figure 3 Spectrogram and pitch contours (top), intensity contours (mid), and annotation panel (bottom) for a sample word tɐ-tɐ HL 'to chew' by a female speaker

An example of a sample sentence with different annotation levels, from which acoustic measurements were taken, is provided in Figure 4.

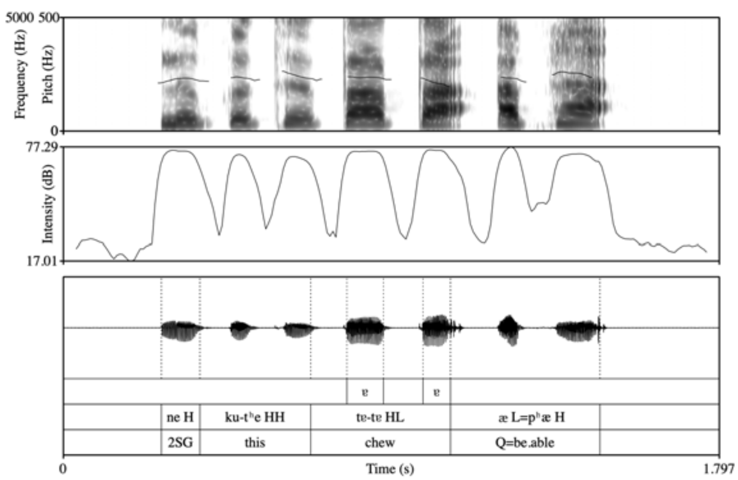


Figure 4 Spectrogram and pitch contours (top), intensity contours (mid), and different annotation levels (bottom) for a sample sentence ne H kutʰe HH tɐ-tɐ HL æ L=pʰæ HL '[Addressing an old person: 'Do you have good teeth,] can you chew this [it's not too hard for you]?' by a female speaker

After manual segmentation of the target words, all measurements were conducted using an automated Praat script. 4234 tokens were analyzed (8 speakers * 32 words * 2 vowels * 3 to 5 repetitions * 2 contexts). The following measurements were made for each vowel:

(i) duration (in ms)

(ii) mean intensity (in dB)

(iii) f0 (in Hz)

The number of measurements varied per parameter, one for duration and mean intensity vs. three for f0: we extracted f0 values at the beginning (1/3), mid-point (2/3), and end of the vowel (3/3) so as to reflect differences in both pitch height and pitch contour (cf., Ladd & Silverman 1984; Johnson et al. 1993). Given cross-gender differences in mean f0 and f0 range (see Wu & Childers 1991; Childers & Wu 1991 for an overview), female and male speakers were analyzed separately.

3.3. Statistical analysis

After segmentation and extraction of measurements, resulting data were subjected to statistical analysis using linear mixed models (Baayen, Davidson, & Bates 2008), implemented in the *lmer* function in the *lme4* package (https://cran.r-project.org/web/packages/lme4/lme4.pdf, Bates, Maechler, Bolker, & Walker 2015) in R (R Development Core Team 2020). The acoustic measurements of duration, intensity, and f0 were included as dependent variables. The fixed effects, along with their levels, were:

- *pitch pattern*: HL, HH, LH
- *syllable position*: 1st, 2nd syllable of the word
- *context*: citation, phrasal (phrase-medial)
- all relevant interactions

The model for f0 had an additional fixed effect *time*, which reflects changes of f0 at three different points over the duration of the vowel.

The random effects included in all models were *speakers* and *words*. Random intercepts and random slopes for speakers and words were included as maximally as permitted by the data (Barr, Levy, Scheepers, & Tily 2013). Maximal models were selected by assessing their Akaike information criterion (AIC) scores, dropping interactions before dropping main effects. Main effects were not dropped, if they contributed to a significant interaction. For post-hoc pairwise comparisons, Bonferroni tests were conducted to compare the respective effects of the three pitch patterns on the first and the second syllable of the word.

4. Results

4.1. Summary of acoustic results

Figures 5-7 present a summary of acoustic results for duration (Figure 5), intensity (Figure 6), and f0 (Figure 7) in the two contexts examined. They reveal considerable differences between the two contexts (citation and phrasal), and between the three pitch patterns in each context. Summarized separately for duration and intensity, on the one hand, and f0, on the other hand, these differences are as follows.[13]

Duration and intensity

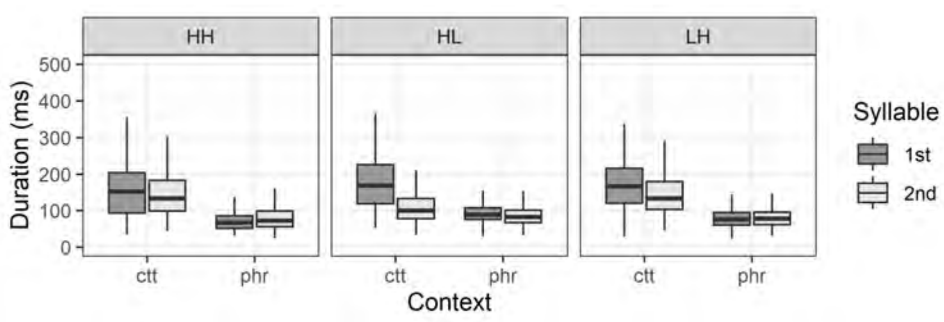


Figure 5 Mean vowel duration in the first and second syllable of the word in the three pitch patterns, averaged over all speakers, in citation form ("ctt") and in phrasal context ("phr")

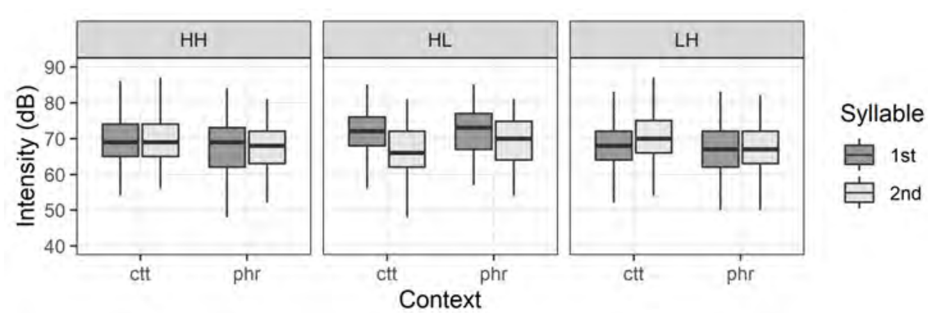


Figure 6 Mean intensity in the first and second syllable of the word in the three pitch patterns, averaged over all speakers, in citation form ("ctt") and in phrasal context ("phr")

As shown in Figures 5-6, in citation context, there appears to be a potentially significant difference in duration and intensity between the first and the second syllable of the word in the pitch patterns HL and LH. Specifically, the pitch pattern HL has longer duration and higher intensity on the first syllable, whereas the pitch pattern LH has longer duration on the first syllable, but higher intensity on the second syllable. By contrast, the pitch pattern HH displays marginal differences between the two syllables of the word, with the first syllable being marginally longer and having marginally higher intensity than the second syllable.

In phrase-medial position, durational differences between the two syllables of the word appear insignificant in all pitch patterns. On the other hand, intensity differences appear insignificant in the pitch patterns HH and LH, but potentially significant in the pitch pattern HL.

F0

As shown in Figure 7, f0 trends are similar for both female and male speakers. The f0 moving trends over the first syllable are broadly similar in all pitch patterns and in both contexts (level or slightly rising). By contrast, f0 over the second syllable shows greater fluctuations in citation context (falling contour in all three pitch patterns) than in phrase-medial position (level tone in the pitch patterns HH and LH, and a less steep falling contour in the pitch pattern HL). In phrase-medial position, the pitch patterns HH and LH show broadly comparable f0 values over the two syllables of the word. By contrast, the pitch pattern HL maintains a clear difference in f0 values between the two syllables, with higher f0 values on the first syllable and lower f0 values on the second syllable, as in citation context.

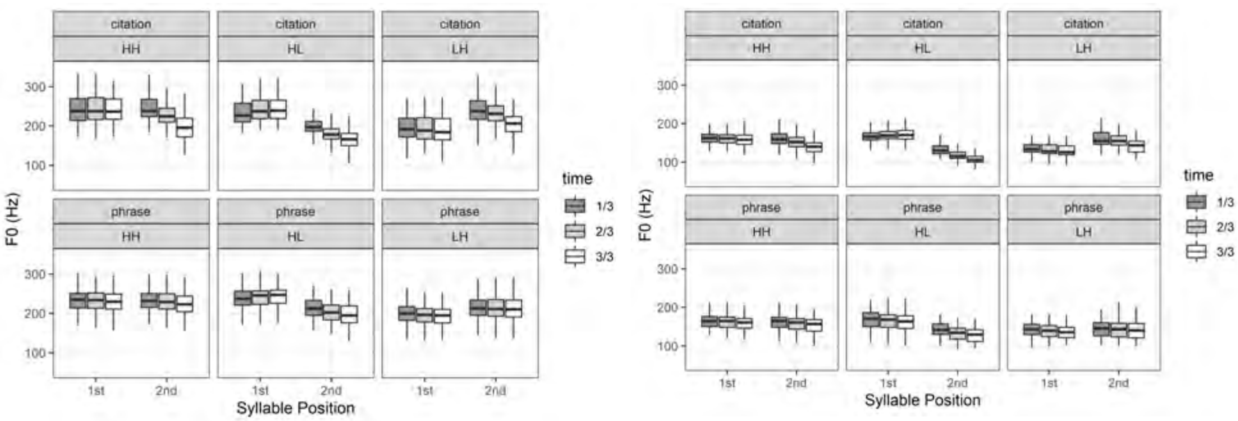


Figure 7 F0 at the 1/3, 2/3, and 3/3 of the vowel duration in the first and second syllable of the word in the three pitch patterns in citation form (top) and in phrase-medial position (bottom), averaged over female speakers (left) and over male speakers (right).

4.2. Results of statistical analysis

This section presents results for the dependent variables duration, intensity, and f0. For each of these dependent variables, we first report results of linear mixed modeling to test main effects of each independent variable (that is, pitch pattern, syllable position, context, and for f0, also time) and the interactions between the independent variables (see section 3.3). Then, in order to explore the results in greater detail for asymmetry in duration and intensity between the two syllables of the word, for each pitch pattern (HL, HH, LH) and each context (citation, phrase) we perform multiple comparisons on syllable position using Bonferroni method. Finally, in order to explore the results for possible reduction of tonal contrasts on either syllable of the word in both contexts examined, we also perform multiple comparisons on the three pitch patterns using Bonferroni method.

4.2.1. Duration

The results of linear mixed modeling for duration in Table 5 reveal significant effects of all independent variables and all of their interactions.

Table 5 Estimates of fixed effects on duration (R^2=.48). Reference categories: PitchPattern#HL, Syl#2, Context#Citation

	Estimate (ms)	SE	df	t value	Pr(>|t|)
(Intercept)	118.18	8.77	12.04	13.47	<.001
PitchPattern#HH	32.06	4.87	1893.32	6.58	<.001
PitchPattern#LH	27.32	4.66	2398.55	5.86	<.001
Syl#1	59.84	4.23	4180.61	14.16	<.001
Context#Phrase	-34.86	5.54	1988.97	-6.3	<.001
PitchPattern#HH*Syl#1	-54.50	5.45	4180.61	-10	<.001
PitchPattern#LH*Syl#1	-34.74	5.33	4180.61	-6.52	<.001
PitchPattern#HH*Context#Phrase	-23.89	7.84	1262.61	-3.05	0.002
PitchPattern#LH*Context#Phrase	-31.10	6.58	2828.09	-4.73	<.001
Syl#1*Context#Phrase	-54.08	6.18	4180.61	-8.75	<.001
PitchPattern#HH*Syl#1*Context#Phrase	36.97	7.88	4180.61	4.69	<.001
PitchPattern#LH*Syl#1*Context#Phrase	22.74	7.71	4180.61	2.95	0.003

The results of pairwise comparisons for duration on syllable position in Table 6 suggest that in citation context, there is a significant difference in duration between the first and the second syllable of the word in the pitch patterns HL and LH (as can also be observed in Figure 5). In both cases, the first syllable is longer than the second syllable (by 59.84 ms in the pitch pattern HL and by 25.10 ms in the pitch pattern LH). On the other hand, in the pitch pattern HH neither syllable is significantly longer than the other (while the first syllable is slightly longer than the second syllable, by 5.35 ms). In phrasal context, the pattern is reversed. The difference between the two syllables in duration is reaching significance in the pitch pattern HH, where the second syllable is longer than the first syllable (by 11.77 ms). There is no significant difference in duration in the pitch pattern HL (where the first syllable is slightly longer than the second syllable), whereas in the pitch pattern LH the difference is marginal (the second syllable is slightly longer than the first syllable).

Table 6 Pairwise comparisons for duration between the 1st and the 2nd syllable by the two contexts (citation, phrase) and three pitch patterns (HL, HH, LH) using Bonferroni method

Context	Pitch Pattern	Comparison I	J	Mean Difference (I-J)	SE	p-value
citation	HL	Syl#2	Syl#1	-59.84	4.23	<.0001
	HH	Syl#2	Syl#1	-5.35	3.44	0.1203
	LH	Syl#2	Syl#1	-25.10	3.25	<.0001
phrase	HL	Syl#2	Syl#1	-5.76	4.51	0.2016
	HH	Syl#2	Syl#1	11.77	3.49	0.0007
	LH	Syl#2	Syl#1	6.24	3.26	0.0556

4.2.2. Intensity

The results of linear mixed modeling for intensity in Table 7 also reveal significant effects of all independent variables and all of their interactions.

Table 7 Estimates of fixed effects on intensity (R^2=.69). Reference categories: PitchPattern#HL, Syl#2, Context#Citation

	Estimate (dB)	SE	df	t value	Pr(>|t|)
(Intercept)	67.20	1.94	7.66	34.59	<.001
PitchPattern#HH	2.84	0.41	3054.21	6.93	<.001
PitchPattern#LH	1.80	0.39	3479.18	4.61	<.001
Syl#1	4.60	0.35	4181.39	13.26	<.001
Context#Phrase	2.17	0.47	3165.52	4.66	<.001
PitchPattern#HH*Syl#1	-4.25	0.45	4181.39	-9.50	<.001
PitchPattern#LH*Syl#1	-7.00	0.44	4181.39	-16.01	<.001
PitchPattern#HH*Context#Phrase	-5.20	0.67	2461.07	-7.82	<.001
PitchPattern#LH*Context#Phrase	-4.46	0.55	3729.61	-8.13	<.001
Syl#1*Context#Phrase	-1.55	0.51	4181.39	-3.06	0.002
PitchPattern#HH*Syl#1*Context#Phrase	1.67	0.65	4181.39	2.59	0.010
PitchPattern#LH*Syl#1*Context#Phrase	3.89	0.63	4181.39	6.16	<.001

The results of pairwise comparisons for intensity on syllable position in Table 8 show that in citation context, there is a significant difference in intensity between the first and second syllable of the word in the pitch patterns HL and LH. In the pitch pattern HL, the first syllable has higher intensity (by 4.6 dB), whereas in the pitch pattern LH, the second syllable has higher intensity (by 2.4 dB) (as can also be observed in Figure 6). By contrast, neither syllable in the pitch pattern HH has significantly higher intensity than the other syllable. In phrasal context, on the other hand, there is a significant difference in intensity between the two syllables of the word in the pitch pattern HL (where the first syllable has higher intensity than the second syllable, by 3.05 dB), while neither syllable in the pitch patterns HH and LH appears to have significantly higher intensity than the other syllable.

Table 8 Pairwise comparisons for intensity between the 1st and the 2nd syllable in the two contexts (citation, phrase) and three pitch patterns (HL, HH, LH) using Bonferroni method

Context	Pitch Pattern	Comparison I	J	Mean Difference (I-J)	SE	p-value
citation	HL	Syl#2	Syl#1	-4.60	0.35	<.0001
	HH	Syl#2	Syl#1	-0.35	0.28	0.214
	LH	Syl#2	Syl#1	2.40	0.27	<.0001
phrase	HL	Syl#2	Syl#1	-3.05	0.37	<.0001
	HH	Syl#2	Syl#1	-0.47	0.29	0.098
	LH	Syl#2	Syl#1	0.06	0.27	0.822

4.2.3. F0

The results of linear mixed modeling for f0 in Table 9 reveal significant effects of all independent variables (pitch pattern, syllable position, context, time) and most of their interactions. Both female and male speakers show similar patterns.

Table 9 Estimates of fixed effects on f0 for male and female speakers (female speakers: R^2=.42, male speakers: R^2=.56). Reference categories: PitchPattern#HL, Syl#2, Context#Citation, Time#1/3

	Female speakers					Male speakers				
	Estimate (Hz)	SE	df	t value	Pr(>|t|)	Estimate (Hz)	SE	df	t value	Pr(>|t|)
(Intercept)	196.18	8.128	4.39	24.14	<.001	130.30	5.40	3.99	24.15	<.001
PitchPattern#HH	51.60	3.896	5841.21	13.24	<.001	33.10	2.12	5613.12	15.60	<.001
PitchPattern#LH	43.80	3.73	6061.26	11.75	<.001	32.33	2.07	5866.84	15.60	<.001
Syl#1	35.51	3.88	6312.74	9.14	<.001	33.96	2.06	6122.32	16.46	<.001
Context#Phrase	13.94	4.46	5792.85	3.13	0.002	14.88	2.38	5727.74	6.24	<.001
Time#2/3	-16.03	3.90	6312.71	-4.11	<.001	-13.41	2.08	6122.24	-6.45	<.001
Time#3/3	-20.00	3.97	6312.87	-5.04	<.001	-23.68	2.13	6122.38	-11.12	<.001
PitchPattern#HH*Syl#1	-39.14	5.02	6312.72	-7.80	<.001	-32.09	2.64	6122.29	-12.16	<.001
PitchPattern#LH*Syl#1	-80.51	4.85	6312.72	-16.61	<.001	-60.19	2.62	6122.29	-22.96	<.001
PitchPattern#HH*Context#Phrase	-30.82	5.98	4955.04	-5.16	<.001	-19.91	3.24	5074.27	-6.14	<.001
PitchPattern#LH*Context#Phrase	-34.93	5.36	6161.90	-6.52	<.001	-30.93	2.91	6035.09	-10.62	<.001
Syl#1*Context#Phrase	-9.56	5.72	6312.71	-1.67	0.095	-8.44	2.98	6122.28	-2.83	0.005
PitchPattern#HH*Time#2/3	-0.88	5.03	6312.70	-0.18	0.861	4.58	2.65	6122.23	1.73	0.084
PitchPattern#LH*Time#2/3	7.89	4.86	6312.71	1.62	0.104	8.58	2.63	6122.23	3.26	0.001
PitchPattern#HH*Time#3/3	-8.27	5.09	6312.81	-1.62	0.104	0.95	2.70	6122.32	0.35	0.724
PitchPattern#LH*Time#3/3	-10.14	4.92	6312.81	-2.06	0.039	5.85	2.67	6122.33	2.19	0.029
Syl#1*Time#2/3	25.59	5.49	6312.70	4.66	<.001	17.62	2.92	6122.23	6.04	<.001
Syl#1*Time#3/3	31.63	5.54	6312.78	5.71	<.001	28.50	2.95	6122.31	9.66	<.001
Context#Phrase*Time#2/3	5.65	5.73	6312.70	0.99	0.324	3.50	2.99	6122.24	1.17	0.242
Context#Phrase*Time#3/3	2.17	5.79	6312.79	0.38	0.708	8.30	3.03	6122.32	2.74	0.006
PitchPattern#HH*Syl#1*Context#Phrase	13.83	7.26	6312.71	1.91	0.057	9.22	3.82	6122.26	2.41	0.016
PitchPattern#LH*Syl#1*Context#Phrase	40.44	7.04	6312.71	5.75	<.001	33.27	3.77	6122.27	8.83	<.001
PitchPattern#HH*Syl#1*Time#2/3	-6.73	7.10	6312.70	-0.95	0.343	-9.83	3.73	6122.23	-2.63	0.008
PitchPattern#LH*Syl#1*Time#2/3	-20.69	6.85	6312.70	-3.02	0.003	-16.56	3.70	6122.23	-4.47	<.001
PitchPattern#HH*Syl#1*Time#3/3	-2.98	7.14	6312.75	-0.42	0.677	-10.38	3.76	6122.28	-2.76	0.006
PitchPattern#LH*Syl#1*Time#3/3	-7.47	6.90	6312.75	-1.08	0.278	-17.36	3.74	6122.28	-4.65	<.001
PitchPattern#HH*Context#Phrase*Time#2/3	7.66	7.27	6312.70	1.05	0.292	1.77	3.83	6122.24	0.46	0.644
PitchPattern#LH*Context#Phrase*Time#2/3	1.56	7.04	6312.70	0.22	0.824	-0.63	3.78	6122.24	-0.17	0.868
PitchPattern#HH*Context#Phrase*Time#3/3	17.12	7.32	6312.76	2.34	0.019	6.61	3.87	6122.29	1.71	0.088
PitchPattern#LH*Context#Phrase*Time#3/3	26.40	7.10	6312.75	3.72	0.0002	5.11	3.81	6122.30	1.34	0.180
Syl#1*Context#Phrase*Time#2/3	-11.99	8.08	6312.69	-1.48	0.138	-10.40	4.21	6122.24	-2.47	0.014
Syl#1*Context#Phrase*Time#3/3	-10.00	8.13	6312.74	-1.23	0.218	-18.68	4.24	6122.27	-4.40	<.001
PitchPattern#HH*Syl#1*Context#Phrase*Time#2/3	-3.74	10.26	6312.69	-0.36	0.716	4.31	5.40	6122.23	0.80	0.425
PitchPattern#LH*Syl#1*Context#Phrase*Time#2/3	4.011	9.95	6312.69	0.40	0.687	7.66	5.32	6122.23	1.44	0.150
PitchPattern#HH*Syl#1*Context#Phrase*Time#3/3	-12.96	10.30	6312.72	-1.26	0.209	3.03	5.43	6122.26	0.56	0.577

PitchPattern#LH*Syl#1*Context# Phrase*Time#3/3	-19.75	9.99	6312.72	-1.98	0.048	4.69	5.35	6122.26	0.88	0.381

As shown in Table 9, 1st syllable position yields higher mean f0 values (estimates are significantly positive), whereas time at the 2/3 and 3/3 duration of the vowel yields lower mean f0 values (that is, mean f0 decreases throughout the syllable duration; estimates are significantly negative). Insignificant interactions for both female and male speakers mainly include interactions of the first syllable in the pitch patterns HH and LH with time in phrasal context (e.g., *Syl#1*Context#Phrase*, *PitchPattern#HH *Time, PitchPattern#HH *Syl#1* Context#Phrase*Time, PitchPattern#LH*Syl#1*Context#Phrase*Time#2/3*). This suggests that in in the pitch patterns HH and LH in phrasal context there is not much difference in mean f0 in the first syllable (as can also be observed in Figure 7).

Both female and male speakers also show similar trends in pairwise comparisons on pitch patterns, as detailed in Table 10. Overall, differences in f0 between the pitch patterns are larger in citation context (see mean difference values in the two contexts in Table 10; differences in f0 pitch range between the two contexts can also be observed in Figure 7). Pairwise comparisons in Table 10 show that for both female and male speakers, in the first syllable of the word in citation context, there is no statistically significant difference between the pitch patterns HL and HH, both pitch patterns being statistically distinct from the pitch pattern LH. Also, for both female and male speakers, in the second syllable in citation context, there is no statistically significant difference between the pitch patterns HH and LH, both pitch patterns being statistically distinct from the pitch pattern HL. In phrasal context we observe some differences between female and male speakers. For female speakers, the observed patterns are similar to those in citation context, that is, no significant difference between the pitch patterns HL and HH in the first syllable and no significant difference between the pitch patterns HH and LH in the second syllable. In addition, the contrast between the pitch patterns HL and LH is also neutralized at the beginning of the second syllable. For male speakers, on the other hand, we rather observe significant differences in f0 between all three pitch patterns in both syllables of the word (with the exception of the contrast between the pitch patterns HL and LH at the beginning of the second syllable). Possible reasons for the variation between female and male speakers in phrasal context include (i) the marginal nature of the contrasts between the pitch patterns HL-HH and HH-LH in female speakers; (ii) individual variation among female speakers (but see endnote 13); and (iii) the restricted size of the data set used in our study. This issue requires further research.

In sum, across the two contexts we observe a contrast between the pitch patterns HH and LH, on the one hand, and the pitch pattern HL, on the other hand. The former pitch patterns show asymmetry in the realization of tone contrasts between the two syllables of the word, being consistenly contrastive in the first syllable of the word, but neutralizing tone contrasts in the second syllable of the word (with some variation between female and male speakers in phrasal context, see above). The latter pitch pattern (HL) stands in a more complex relationship with the remaining two pitch patterns. It is not statistically distinct from the pitch pattern HH in the first syllable (with some variation between female and male speakers in phrasal context), but it does contrast with the pitch pattern HH in the second syllable. The pitch pattern HL also shows a tendency towards neutralization of contrast with the pitch pattern LH in the second syllable in phrasal context.

Table 10 Pairwise comparisons for f0 on pitch patterns (HL, HH, LH) for female and male speakers in the two contexts (citation, phrase) and the two syllable positions using Bonferroni method. Pitch patterns that are not significantly distinct (p<.0001) are highlighted in gray

					F0 at 1/3			F0 at 2/3			F0 at 3/3		
Gender	Context	Pitch Pattern	Comparison		Mean Diff.	SE	p-value	Mean Diff.	SE	p-value	Mean Diff.	SE	p-value
			I	J	(I-J)			(I-J)			(I-J)		
female	citation	Syl#1	HL	HH	-12.45	3.89	0.0041	-4.84	3.89	0.6393	-1.21	3.89	1.0000

			HL	LH	36.71	3.72	<.0001	49.51	3.72	<.0001	54.32	3.72	<.0001
			HH	LH	49.16	3.32	<.0001	54.35	3.32	<.0001	55.53	3.32	<.0001
		Syl#2	HL	HH	-51.60	3.9	<.0001	-50.71	3.9	<.0001	-43.33	3.99	<.0001
			HL	LH	-43.80	3.73	<.0001	-51.70	3.73	<.0001	-33.67	3.81	<.0001
			HH	LH	7.79	3.32	0.0567	-0.98	3.32	1.0000	9.66	3.34	0.0114
	phrase	Syl#1	HL	HH	4.54	4.1	0.8056	8.22	4.1	0.1353	11.62	4.1	0.0139
			HL	LH	31.20	3.94	<.0001	38.42	3.94	<.0001	42.17	3.94	<.0001
			HH	LH	26.66	3.29	<.0001	30.20	3.29	<.0001	30.55	3.29	<.0001
		Syl#2	HL	HH	-20.77	4.11	<.0001	-27.56	4.11	<.0001	-29.62	4.13	<.0001
			HL	LH	-8.87	3.94	0.0736	-18.33	3.94	<.0001	-25.13	3.96	<.0001
			HH	LH	11.90	3.3	0.0009	9.23	3.3	0.0154	4.49	3.3	0.5197
male	citation	Syl#1	HL	HH	-1.01	2.1	1.0000	4.23	2.1	0.1329	8.42	2.1	0.0002
			HL	LH	27.86	2.06	<.0001	35.84	2.06	<.0001	39.37	2.06	<.0001
			HH	LH	28.87	1.78	<.0001	31.6	1.78	<.0001	30.95	1.78	<.0001
		Syl#2	HL	HH	-33.1	2.13	<.0001	-37.68	2.12	<.0001	-34.05	2.17	<.0001
			HL	LH	-32.33	2.07	<.0001	-40.91	2.07	<.0001	-38.18	2.13	<.0001
			HH	LH	0.78	1.78	1.0000	-3.23	1.78	0.2087	-4.12	1.79	0.0642
	phrase	Syl#1	HL	HH	9.68	2.18	<.0001	8.85	2.17	0.0001	9.48	2.17	<.0001
			HL	LH	25.52	2.08	<.0001	26.46	2.08	<.0001	27.22	2.08	<.0001
			HH	LH	15.84	1.85	<.0001	17.62	1.85	<.0001	17.75	1.85	<.0001
		Syl#2	HL	HH	-13.19	2.18	<.0001	-19.54	2.18	<.0001	-20.75	2.19	<.0001
			HL	LH	-1.4	2.09	1.0000	-9.35	2.08	<.0001	-12.36	2.09	<.0001
			HH	LH	11.8	1.86	<.0001	10.19	1.86	<.0001	8.39	1.86	<.0001

5. Discussion

In this section, the results in Section 4 are discussed in relation to the research goals and predictions in section 2.3.

In relation to our main goal—to empirically examine the previous claims of stress in Kala Lizu in Chirkova & Chen (2013)—our acoustic results and statistical analyses are consistent with observations in that study in relation to words in citation form. Specifically, in citation context, there is evidence for asymmetry in duration and intensity in the pitch patterns HL and LH, and no similar asymmetry in the pitch pattern HH.

Our acoustic results and statistical analyses also reveal asymmetries in intensity and, marginally, duration in phrasal context. The observed patterns are, however, distinct from those in citation context. Specifically, in phrasal context (i) there is asymmetry in intensity in the pitch pattern HL, and no asymmetry in intensity in the pitch patterns HH and LH; and (ii) there is marginal evidence for asymmetry in duration in the pitch patterns HH and LH, and no asymmetry in duration in the pitch pattern HL. In addition, across both citation and phrasal contexts, there is evidence for asymmetry in the realization of tone contrasts between the two syllables of the word in the pitch patterns HH and LH. Overall, these findings confirm that Kala Lizu has contrastive differences in the relative prominence of syllables in disyllabic words in both contexts examined.

On the whole, the observed patterns are complex, but they do allow for a uniform interpretation across the two contexts, phrasal and citation, using evidence for asymmetries in intensity and those in the realization of tone contrasts. Durational differences between the two syllables of the word are marginal and do not appear to meaningfully correlate with the observed asymmetries in intensity and in the realization of tone contrasts. For that reason, they are not included in the discussion below.

(1) The phrasal context, which in our study is taken to provide more straightforward evidence on word-level stress, can be tentatively analyzed as follows.

Assuming that intensity is associated with word-level stress, the pattern HL can be interpreted as having stress-like characteristics. Higher intensity on the first syllable and reduced intensity on the second syllable suggests the first syllable as the position of prominence in that pitch pattern. The association of stress with intensity is in line with the tendency for tonal languages to employ intensity as a marker of stress (e.g. Papiamentu, Remijsen & van Heuven 2005; see Gordon & Roettger 2017: 5–6).

The two pitch patterns that do not show asymmetry in intensity (HH and LH) can be further analyzed by taking f0 as the main acoustic correlate of tone. These two pitch patterns have the following characteristics. They are contrastive on the first syllable of the word. The contrast between H(H) and L(H) (on the first syllable of the word) corresponds to the full set of tonal contrasts on monosyllabic words (H vs. L). The pitch patterns HH and LH on disyllabic words also correspond to the two pitch patterns that can be regularly derived from the two contrastive lexical tones on monosyllabic words in nominal compounds (see the first sandhi rule in Section 2.1). Therefore, we analyze these two pitch patterns as having tone-like characteristics. The contrast between the pitch patterns HH and LH is restricted to the first syllable, and tends to be neutralized over the second syllable. The full realization of the tones on the first syllable and their neutralization on the second syllable can be interpreted as a strong-weak stress-like pattern, in which the first syllable, characterized by retention of the full set of tonal contrasts, is the position of prominence. Tonal contrasts are reduced in the weak, unstressed position. Tone placement in these pitch patterns can therefore be analyzed as stress dependent.

In sum, across the three pitch patterns in phrasal context, the first syllable of the word can be seen as the position of prominence, albeit associated with different acoustic properties (intensity in the case of the pitch pattern HL and full tone realization in the case of the pitch patterns HH and LH). This suggests that disyllabic words in Kala Lizu have fixed stress in word-initial position.

(2) The citation context, where the word constitutes an utterance of its own and is phrase-final, displays greater intensity, and also longer duration and a broader f0 range associated with the first syllable (see Figures 5-7). (Note that the pitch pattern LH is exceptional in that greater intensity in that pattern is found on the second syllable, see below). A combination of greater intensity, longer duration, and a broader f0 range on the first syllable can be interpreted as signaling phrasal (nuclear) stress.[14] That further entails that phrase-level stress in Kala Lizu occurs on stressed syllables (within a word), as common in many languages with word-level stress.

In relation to boundary tones, we can assume one boundary tone for all three pitch patterns, that is, HL% (as also found on monosyllabic words, see Section 2.1). In the pitch pattern LH, the addition of the boundary HL% tone is associated with an increase in intensity. This may be due to the tendency for f0 and intensity to co-vary during speech production (e.g. Hirano, Ohala & Vennard 1969; Tilsen 2016).

Our analysis of pitch patterns on disyllabic words in Kala Lizu as having fixed word-initial stress, and combining one stress-like pattern (HL) and two tone-like patterns (HH and LH) contributes to a better understanding of the prosodic organization of that language. Notably, it offers an explanation for the lack of minimal three-way lexical contrasts between the three pitch patterns. That lack is due to the binary contrast between prosodic word types in this language: more stress-like and more tone-like. Possible contrasts between words hence include: (i) those between words with the stress-like pattern and words with the tone-like patterns (i.e., HL vs. HH, HL vs. LH), and (ii) those between words with the two tone-like patterns (i.e., HH vs. LH).

Seen in the context of other Lizu varieties, our analysis of pitch patterns on disyllabic words in Kala Lizu also sheds light on complex correspondences illustrated by Table 1. Specifically, all Lizu varieties share the tone-like patterns HH and LH (which are likely connected to the two lexical tones on monosyllabic words, /H/ and /L/). Put differently, all varieties appear to share lexical tone, which may be combined with stress (as in Kala Lizu, as discussed presently). In addition to the two tone-like patterns, shared by all varieties, some varieties also contain different proportions of words with the stress-like pattern (HL) (Kala Lizu 1, Kala Lizu 2 in Table 1). Since hierarchies of borrowability typically rank prosody at the top of adopted features (e.g., Thomason & Kaufman 1988, 75; Salmons 1992; Matras 2009, 231–233), developments affecting pitch patterns on polysyllabic domains in Lizu are likely susceptible to restructuring in a language-contact setting. In the case of Kala Lizu, discussed in this study, the hybrid prosodic organization (with lexical tone and fixed, word-initial stress) has clear parallels in the language with which Lizu is historically in contact: Tibetan. Tonal Tibetan dialects spoken in China mostly have simple tonal inventories, consisting of two basic lexical tones: High (H) versus Low (L). In addition, tonal Tibetan dialects are traditionally analyzed as having word-initial stress (e.g., Geziben 1996; Sun 1997, 2003). Arguments for the existence of stress include (i) stress-like vowel quality alternations in initial syllables as opposed to non-initial syllables in polysyllabic words, and (ii) constraints on pitch patterns in polysyllabic words, whereby the tone on the first syllable in a disyllabic word can be either H or L, but the tone on the second syllable can only be H. A combination of lexical tone (with the basic contrast between H and L) with fixed, word-initial stress (tone being dependent on stress, with tone reduction in non-stressed syllables), as described in relation to our Kala Lizu data, is therefore similar to that found in tonal Tibetan dialects. The development in Kala Lizu of the stress-like pattern HL, which occurs with a similar frequency as the two tone-like patterns, requires further study. The development of that pattern can be tentatively correlated with the development of derivational verbal morphology, most importantly directional prefixes, which are lexically unspecified for tone. Some tonal Tibetan dialects of Sichuan in contact with local non-Tibetic languages with developed verbal morphology (Qiangic languages) provide examples of how such a development might happen. Note that the Tibetan dialects that acquire directional prefixes (commonly grammaticalized from imperative prefixes, such as WT *ya* 'up', *ma* 'down') also show an additional pitch pattern that is not correlated with the two basic lexical tones H and L (i.e., HL, see Chirkova 2014 for Kami Tibetan).

It is possible that tendencies in stress location and the acoustic encoding of stress in other varieties of Lizu may replicate the patterns of their respective contact languages.[15] These include, in addition to Tibetan and Mandarin, languages with dense tonal specification and possibly, without stress (such as Nuosu Yi, Edmondson, Esling, and Lama 2017), and languages with sparse tonal specification and possibly, stress (such as Pumi, e.g., Greif 2010). The methodology adopted in this study suggests itself as a promising way for the exploration of the prosodic systems of other Lizu varieties, and also of their contact languages.

6. Conclusion

In this study we presented an acoustic investigation of the three surface pitch patterns on disyllabic words in Kala Lizu. Using the parameters of duration, intensity, and f0 we attempted to sort out different forms of prominence in that language, and to explore their interaction with lexical tone. The acoustic results and statistical analyses indicate that the first syllable in Kala Lizu is the position of prominence in all three pitch patterns, suggesting that Kala Lizu has fixed stress in word-initial position. Our data further reveal a binary distinction among the three surface pitch patterns: between the pitch pattern HL, which may be interpreted as more stress-like, on the one hand, and the pitch patterns HH and LH, which may be interpreted as more tone-like, on the other hand. In the pitch pattern HL, intensity is an acoustic correlate of stress; whereas in the pitch patterns HH and LH, stress is cued by full realization of tones, whereas the tone on the second syllable may be neutralized. Based on these findings, we propose that Kala Lizu has a hybrid prosodic system combining tone and word-level stress, where tone is dependent on stress.

Naturally, further work and additional data from more speakers are required to test the robustness of present conclusions. In particular, there is a need for conducting perception experiments, for in their absence, the reported distinctions may not reflect true differences for Lizu speakers. It would also be crucial to study corpora that are more balanced for word type and, for that reason are more conclusive as to the potential impact of morphosyntactic structures on the three pitch patterns or the lack thereof (as assumed in the present study). The sample size and the number of items in experimental paradigms also need to be increased, so as to enhance statistical power in data analysis and to get a better understanding of cross-gender differences observed in phrasal context.

The results of our study are admittedly preliminary and require further research. At the same time, they open up a host of possibilities for gaining a better understanding of synchronic variation in the prosodic organization of Lizu varieties as well as general principles governing the prosodic system of the Lizu language.

Appendix. The dataset.

Word list

HL	**HH**	**LH**
be-be ‘to climb’		be-be ‘to crawl’
		bu-bu ‘to explode’
dɐ-dɐ ‘to be short’		
tɐ-tɐ ‘to chew’	tɐ-tɐ ‘to hold in one’s arms’	jɐ-tɐ ‘left’
	tsʰæ-tsʰæ ‘to be hot’	
tɕy-tɕy ‘to be straight’	tsʰu-tsʰu ‘to be thick’	
		tʃʰɐ-tʃʰɐ ‘magpie’
		tʃʰu-tʃʰu ‘mouth harp’
	Ndzu-Ndzu ‘button’	dzu-dzu ‘to be steep’
kæ-kæ ‘to fiddle’	kæ-kæ ‘to fight’	
gæ-gæ ‘to amuse oneself’		
	kælæ ‘hardship’	kælæ ‘butterfly’
	kʰæ-kʰæ ‘to split, divide’	
qɐ-qɐ ‘to filter’	qɐtɐ ‘knot’	qwɐ-qwɐ ‘to be hard’
sɐ-sɐ ‘to touch, stroke’		
		zu-zû ‘thread (needle)’
		zæ-zæ̂ ‘be tender’
	ʃo-ʃo ‘to be clean’	ʃɐ-ʃɐ ‘to be long’
		ʃæ-ʃæ ‘to search’
	ʒu-ʒu ‘to be narrow’	ʒæ-ʒæ ‘to quarrel’
	ɲy-ɲy ‘to be soft’	

Frame sentences

Frames sentences differ depending on the part of speech of the target word.[16]

(i) Verbs (a total of 14):

ne H/tʰe-ɹə HL [target] æ L=pʰæ HL / mæ H=pʰæ L

2SG/3-PL Q=be.able/NEG=be.able

‘Can you / they [target]? // You / they cannot [target].’

For example:

ne H ku-tʰe HH tɐ-tɐ HL æ L=pʰæ HL?

2SG this-that chew-chew Q=be.able

‘[Addressing an old person] (Do you have good teeth,) can you chew this [it’s not too hard for you]?’

Other common frames include:

tʰe-ɹə HL	[target]	mæ H=tʃu L
3-PL		NEG=be.likely

‘They are not likely to VERB.’

For example:

zuɹə HL	kæ-kæ HH	mæ H=tʃu L
3-PL	fight-fight	NEG=be.likely

‘They are not likely to fight.’

(ii) Adjectives (11 in total):

ku-tʰe HH	NOUN	dɐzu LH	[target]	to HL
this-that		really		VIS

‘This NOUN is really [target].’

For example:

ku-tʰe HH	ɹəpʰæ HL	dɐzu LH	ʒu-ʒu HH	to L,
this-that	road	really	be.narrow-be.narrow	VIS
fu-fu HH	mæ H=lje L			
walk-walk	NEG=be.good			

‘This path is really narrow and difficult to walk.’

(iii) Nouns (7 nouns):

æ H	[target]	VERB=ɐ LH
3SG		VERB= CS

‘I [target]=ed.’

For example:

æ H	kælæ LH	te-pu LH	kʰe-Ndo=ɐ LH
1SG	butterfly	one-CLF	inward-see=CS

'I saw a butterfly.'

Notes

[1] Muli Lizu 1 refers to the variety spoken in Manao 玛瑙村 and Majidian 马鸡店 villages in Kala Township, Muli County. This variety is locally known as ʃætɕʰopæ HHH or 'eastern dialect' (from WT *shar phyogs* 'east'). Data collected by the first author. Muli Lizu 2 refers to a slightly different variety spoken in Kala Township, as reported in Huang & Renzeng (1991) and Dai et al. (1992). Mianning Lizu refers to the variety spoken in Lagusa 拉姑萨 village, He'ai 和爱 Township, Mianning County, data from Yu (2012). Jiulong Lizu refers to the variety spoken in Naiqu village, Naiqu Township, Jiulong County, data from Ikeda (2009). The notation of the tones in various sources (with Chao tone numbers [Chao 1930] for Muli Lizu 2 and Naiqu Lizu; and with a grave accent /`/ for the high tone, the low tone being unmarked for Mianning Lizu in Yu 2012) are converted to H for the high tone and L for the low tone for ease of comparison.

[2] In Yu's (2012) analysis, every prosodic word, regardless of the number of syllables it contains, is either high- or low-toned. High-toned words may have surface pitch patterns HH and HL (these two pitch patterns appear to be in free variation), whereas low-toned words have the surface pitch pattern LH.

[3] Monomorphemic polysyllabic words are also found, but these tend to be loanwords. Examples include: Kala Lizu kwɐtsə HL '(sunflower) seeds', from Southwest Mandarin /kwa⁴⁴tsə⁵³/ 瓜子; Kala Lizu ʃæwu HH 'paper', from Tibetan, WT *shog bu*.

[4] Tone analysis presented in Figure 1 was carried out using scripts developed by James Stanford for the software packages Praat (Boersma & Weenick 2020) and R (R Development Core Team 2020). Stanford's scripts normalize syllable tokens for time duration. The raw input file is converted to 200 "relative time" points, which can be compared using mean pitch values at selected relative time points. The f0 normalization process uses the mean of the Lizu low-rising tone as a point of reference for tone comparison. Tone inventory is presented in semitones, given that semitones better reflect pitch-related perception than the Hertz scale. (For more details see Stanford, 2008, 420–421 and references therein.)

[5] The tendency for high-tone monosyllabic words to be realized with a fall in pitch and for low-tone monosyllabic words to be realized with a rise in pitch, when uttered in isolation, is commonly shared by many linguistic neighbors of Lizu that have simple tone systems. These include Tibetic languages (e.g., J. Sun 1997, 491, 499) and other Qiangic languages (see Evans 2008, 484 for an overview; this latter work also suggests the obligatory presence of a high tone in a phonological word in Qiangic tone systems).

[6] At the same time, native speakers seem to be very sensitive to the three different pitch patterns, as suggested by a pilot perception experiment in Chen & Chirkova (2012). In that experiment, the segmental composition of 10 triplets (of comparable segments but with different pitch patterns) was substituted with the syllable /a/, resynthesized with the same acoustic parameters (intensity, duration, and f0) of each corresponding word. Listeners who were asked to identify the synthesized word they heard as one out of the three choices (HL, HH, LH) showed a high degree of consistency in their judgements.

[7] The same rule applies on words and compounds of over two syllables, which have three pitch patterns, comparable to the ones over disyllabic words and compounds. In the case of words and compounds that begin with a disyllabic root or word, the resulting surface pitch pattern is determined by the pitch pattern of the initial disyllabic item. The pitch pattern of the initial disyllabic item is expanded over the resulting compound, whereas tones or pitch patterns on non-initial syllables are neutralized. Consider examples in Table 3a.

Table 3a Examples of compounds consisting of a disyllabic word followed by a monosyllabic word

	ndʒə L 'skin'	mu H 'animal hair'
bɹəɹə HL 'snake' tombu HL 'nose'	bɹəɹə ndʒə HLL 'snake skin'	 tombu mu HLL 'nose hair'
HH: sə̂ŋgê 'lion'	səŋge ndʒə HHH 'lion skin'	səŋge mu HHH 'lion fur'
LH: mutsə̂ 'cat'	mutsə ndʒə LLH 'cat's skin'	mutsə mu LLH 'cat's fur'

[8] Verbal prefixes in Kala Lizu include: (i) directional prefixes: de- 'upward', ne- 'downward', kʰe- 'inward'; and (ii) a perfectivizing suffix, tʰe-.

[9] This is different from some neighboring languages, where different word formation types may have their own tone patterns or tone sandhi rules. Yongning Na is a case in point, for its reduplicated forms have specific tonal templates (see Michaud 2017, 216, 252–258).

[10] Spectral tilt and formant frequencies are also promising correlates of stress. However, as noted in Gordon & Roettger (2017, 5–6), their reliability relative to other potential correlates of stress is difficult to establish definitively given a relatively limited number of studies exploring these parameters, the diversity of their implementations, and, in relation to formant frequencies, the fact that effects often do not hold for all vowels.

[11] For that reason, an unstressed open vowel may seem more stressed than a closed stressed vowel, as in the English noun ˈ*impact* (van Heuven 2018, 19).

[12] As explained in van Heuven (2018, 24), intensities of speech sounds may vary considerably, for instance, when the speaker inadvertently turns his head. Intensity drops due to such variations are of similar magnitude as those reported as correlates of stress (in the order of 5 dB). For these reasons, a constant distance between the mouth of the speaker and the microphone was an important requirement for our study.

[13] Error bars in Figures 5-7 reveal some interpersonal variation, likely reflecting gender- and/or age-related differences in, among others, speech tempo and loudness. Note that in the statistical part of the study, the impact from individual speakers is mitigated by adding speaker as a random effect.

[14] Our data (words in citation context, where the word is an utterance on its own, being at the same time domain-initial and domain-final, and bearing the main stress of the phrase) do not allow us to consider possible interaction of phrase-level stress with another potentially relevant prosodic factor: domain-initial strengthening (e.g., Fougeron & Keating 1997; Fougeron 2001; Cho 2001, 2016; Cho & Keating 2001, 2009). This issue needs to be explored in future research.

[15] The phonological contact-behavior of Romani dialects, as discussed in Matras (2009: 231), provides an example of how stress location in different varieties of the same language may reflect stress patterns in their various contact languages. The conservative stress pattern in Romani is on the final inflectional segment of the word. However, Romani dialects in Western Europe have adopted a strong tendency toward word-initial stress, while those in a zone in central Europe show a moderate tendency toward penultimate stress, in both cases replicating the patterns of the contact languages.

[16] Abbreviations used in the gloss below follow the Leipzig Glossing Rules (LGR, http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Non-standard abbreviations (those not included in the LGR) are: CS = change of state marker, EGO = egophoric, VIS = visual evidential marker.

References

BAAYEN, R.H., D.J. Davidson, and D.M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390–412.

BARR, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3): 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

BATES, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1): 1–48.

BECKMAN, Mary E., and Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

BOERSMA, Paul, and David Weenink. 2020. Praat: Doing phonetics by computer. http://www.praat.org/.

BOLINGER, Dwight L. 1958. A theory of pitch accent in English. *Word* 14: 109–149.

BRUNELLE, Marc. 2017. Stress and phrasal prominence in tonal languages: The case of Southern Vietnamese. *Journal of the International Phonetic Association* 47(3): 283–320.

BYRD, Dani, Jelena Krivokapić, and Sungbok Lee. 2006. How far, how long: On the temporal scope of phrase boundary effects. *Journal of the Acoustical Society of America* 120: 1589–1599.

CABALLERO, Gabriela, and Lucien Carroll. 2015. Tone and stress in Choguita Rarámuri (Tarahumara) word prosody. *International Journal of American Linguistics* 81(4): 457–493.

CAMPBELL, Nick, and Mary Beckman. 1997. Stress, prominence, and spectral tilt. In *Intonation: Theory, Models and Applications (Proceedings of an ESCA Workshop, September 18-20, 1997, Athens, Greece), edited by*

Antonis Botinis, Georgios Kouroupetroglou, & George Carayiannis, 67–70. ESCA and University of Athens Department of Informatics, Greece.

CANGEMI, Francesco, and Stefan Baumann. 2020. Integrating phonetics and phonology in the study of linguistic prominence. *Journal of Phonetics* 81: 1–6.

CHAO, Yuen Ren. 1930. ə sɪstəm əv "toun-lɛtəz". *Le Maître Phonetique* 30 : 24–27.

CHAVEZ-PEON, Mario. 2008. Phonetic cues to stress in a tonal language: Prosodic prominence in San Lucas Quiavini Zapotec. In *2008 CLA (Canadian Linguistic Association) Proceedings*, edited by S. Jones. Vancouver: University of British Columbia. http://www.chass.utoronto.ca/~cla-acl/actes2008/actes2008.html.

CHEN, Yiya, and Carlos Gussenhoven. 2008. Emphasis and tonal implementation in Mandarin Chinese. *Journal of Phonetics* 36: 724–46.

CHEN, Yiya, and Katia Chirkova. 2012. The prosodic system of Lizu: An experimental probe. Presentation at the 45th International Conference on Sino-Tibetan Languages and Linguistics, Nanyang Technological University, Singapore, October 26-28, 2012.

CHILDERS, D.G. and K. Wu. 1991. Gender recognition from speech. Part II: Fine analysis. *Journal of the Acoustical Society of America* 90: 1841–1856.

CHIRKOVA, Katia. 2014. Kami. In *Phonological profiles of little-studied Tibetic varieties*, edited by Jackson T.-S. Sun, 1–75. Taipei: Academia Sinica, Institute of Linguistics.

CHIRKOVA, Katia, and Yiya Chen. 2013. Lizu: Illustrations of the IPA. *Journal of the International Phonetic Association* 43(1): 75–86.

CHO, Taehong. 2001. The effects of prosody on articulation in English. PhD diss., University of California.

CHO, Taehong. 2016. Prosodic boundary strengthening in the phonetics-prosody interface. *Language and linguistics compass* 10(3): 120–141.

CHO, Taehong, and Patricia Keating. 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155–190.

CHO, Taehong, and Patricia Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37: 466–485.

CROSSWHITE, Katherine. 2004. Vowel reduction. In *Phonetically Based Phonology*, edited by B. Hayes, D. Steriade, and R. Kirchner, 191–231. New York: Cambridge University Press.

DAI, Qingxia 戴庆厦, Huang Bufan 黄布凡, Xu Shouchun 许寿椿, Chen Jiaying 陈嘉瑛, and Wang Huiyin 王会银 , eds. 1992. *Zang-Mian yuzu yuyan cihui* 藏缅语族语言词汇 (A Tibeto-Burman lexicon). Beijing: Zhongyang minzu daxue chubanshe.

DE JONG, Kenneth. 2004. Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics* 32: 493–516.

DE JONG, Kenneth, and Bushra Zawaydeh. 2002. Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics* 30: 53–75.

DOWNING, Laura J. 2010. Accent in African languages. In *A Survey of World Accentual Patterns in the Languages of the World*, edited by Rob Goedemans, Harry van der Hulst, and Ellen van Zanten, 381–427. Berlin: Mouton de Gruyter.

FLETCHER, Janet. 2010. The Prosody of Speech: Timing and Rhythm. In *The Handbook of Phonetic Sciences*, edited by William J. Hardcastle, John Laver, and Fiona E. Gibbon, 523–602. Malden, MA and Oxford: Wiley-Blackwell.

FOUGERON, Cécile. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29: 109–135.

FOUGERON, Cécile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106: 3728–3740.

FRY, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765–768.

FRY, D. B. 1958. Experiments in the perception of stress. *Language and Speech* 1: 120–152.

EDMONDSON, Jerold A., John H. Esling, Lama Ziwo. 2017. Nuosu Yi. *Journal of the International Phonetic Association* 47(1): 87–97.

EVANS, Jonathan. 2008. African tone in the Sinosphere. *Language and Liguistics* 9(3): 463–490.

EVERETT, Keren. 1998. The acoustic correlates of stress in Pirahã. *Journal of Amazonian Languages* 1(2): 104–162.

GEZIBEN, Deji-Sezhen. 1996. Trochaic structure in Tibetan phonology: A metrical analysis of tone in Lhasa Tibetan. PhD diss., University of Texas at Arlington.

GORDON, Matthew. 2011. Stress: Phonotactic and phonetic evidence. In *The Blackwell Companion to Phonology*, edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice, 924–948. Malden, MA and Oxford: Wiley-Blackwell.

GORDON, Matthew. 2014. Disentangling stress and pitch accent: Toward a typology of prominence at different prosodic levels. In *Word stress: Theoretical and Typological issues*, edited by Harry van der Hulst, 83–118. Cambridge: Cambridge University Press.

GORDON, Matthew, and Timo Roettger. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3(1): 1–11.

GREIF, Markus. 2010. Tone and intonation in Pumi – a first survey. *Language Typology and Universals STUF, Akademie Verlag* 63(3): 221–251.

GRUBER, James. 2011. An acoustic, articulatory, and auditory study of Burmese tone. PhD diss., Georgetown University, Washington D.C.

GUION, Susan G., Jonathan D. Amithb, Christopher S. Dotya, and Irina A. Shport. 2010. Word-level prosody in Balsas Nahuatl: The origin, development, and acoustic correlates of tone in a stress accent language. *Journal of Phonetics* 38: 137–166.

GUSSENHOVEN, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

HAYES, Bruce. 1989. The Prosodic Hierarchy in meter. In *Rhythm and Meter*, edited by Paul Kiparsky and Gilbert Youmans, 201–260. Orlando, FL: Academic Press.

HINTZ, Diane M. 2006. Stress in South Conchucos Quechua: A phonetic and phonological study. *International Journal of American Linguistics* 72: 477–521.

HIRANO, Minoru, John Ohala, and William Vennard. 1969. The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *Journal of Speech, Language, and Hearing Research* 12(3): 616.

HUANG Bufan 黄布凡 and Renzeng Wangmu 仁增旺姆. 1991. Lüsuyu 吕苏语 (The Lüsu language). In *Zang-Mianyu shiwu zhong* 藏缅语十五种 (Fifteen Tibeto-Burman Languages), edited by Dai Qingxia 戴庆厦, Huang Bufan 黄布凡, Fu Ailan 付爱兰, Renzeng Wangmu 仁增旺姆, and Liu Juhuang 刘菊黄, 132–152. Beijing: Beijing yanshan chubanshe.

HYMAN, Larry M. 2001. Privative tone in Bantu. In *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics: December 12-14, 2000, Gakushi Kaikan, Tokyo*, edited by Shigeki Kaji, 237–257. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

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

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

JOHNSON, Keith, Peter Ladefoged, and Mona Lindau. 1993. Individual differences in vowel production. *Journal of the Acoustical Society of America* 94: 701–714.

IKEDA Takumi. 2009. 200 Basic words of the Lyuzu language (Naiqu dialect). In *Progressive Report*, edited by Yasuhiko Nagano, vol. 3, Project No. 16102001, Grant-in-Aid for Scientific Research.

KALAND, Constantijn. 2019. Acoustic correlates of word stress in Papuan Malay. *Journal of Phonetics* 74: 55–74.

KLATT, Dennis H. 1975. Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics* 3: 129–140.

KRATOCHVIL, Paul. 1964. Disyllabic stress patterns in Peking dialect. *Archiv Orientální* 32: 383–402.

KRISTOFFERSEN, Gjert. 2000. *The Phonology of Norwegian*. Oxford: Oxford University Press.

LADD, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

LADD, D. Robert and Kim E. A. Silverman. 1984. Intrinsic pitch of vowels in connected speech. *Phonetica* 41: 31–40.

LAHIRI, Aditi, Allison Wetterlin, and Elisabet Jönsson-Steiner. 2005. Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics* 28(1): 61–96.

LEHISTE, Ilse, and Gordon E. Peterson. 1959. Vowel amplitude and phoneme stress in American English. *Journal of the Acoustical Society of America* 31: 428–435.

LIN Maocan 林茂灿, and Yan Jingzhu 颜景助. 1980. Beijinghua qingsheng de shengxue xingzhi 北京话轻声的声学性质 (The acoustic nature of the neutral tone in Beijing Mandarin). *Fangyan* 方言 3: 161–165.

LIN, You-Jing. 2009. Units in Zhuòkējì rGyalrong discourse: prosody and grammar. PhD diss., University of California at Santa Barbara.

LIN, You-Jing. 2012. By no means marginal: Privative tone in Zhuokeji Rgyalrong. *Language and Linguistics* 13(4): 665–662.

MATRAS, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

MICHAUD, Alexis. 2017. *Tone in Yongning Na: Lexical Tones and Morphotonology*. (Studies in Diversity Linguistics 13). Berlin: Language Science Press.

PIKE, Eunice V. & Joy Oram. 1976. Stress and tone in the phonology of Diuxi Mixtec. *Phonetica* 33: 321–333.

PETERSON, Gordon E. and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32: 693–703.

POTISUK, Siripong, Jackson Gandour, and Mary P. Harper. 1996. Acoustic correlates of stress in Thai. *Phonetica* 53: 200–220.

R DEVELOPMENT CORE TEAM. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org, accessed July 2020.

REMIJSEN, Bert. 2002. Lexically contrastive stress accent and lexical tone in Ma'ya. In *Laboratory Phonology 7*, edited by Carlos Gussenhoven, and Natasha Warner, 585–614. Berlin: Mouton de Gruyter.

REMIJSEN, Bert, and Vincent van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* 22: 205–235.

ROETTGER, Timo, and Matthew Gordon. 2017. Methodological issues in the study of word stress correlates. *Linguistic Vanguard* 3.1, 1-11. DOI: 10.1515/lingvan-2017-0006

SALMONS, Joe. 1992. *Accentual Change and Language Contact: Comparative Survey and a Case Study of Early Northern Europe*. London: Routledge.

SHEN, Xiaonan Susan. 1993. Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech* 36: 415–433.

SHIH, Chi-lin. 1988. Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory* 3: 83–109.

SLUIJTER, Agaath M. C., and Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100: 2471–2485.

SLUIJTER, Agaath M. C., Vincent J. van Heuven, and Jos J. A. Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America* 101: 503–513.

SMITH, Caroline, Donna Erickson, and Christophe Savariaux. 2019. Articulatory and acoustic correlates of prominence in French: Comparing L1 and L2 speakers. *Journal of Phonetics* 77: 1–29.

SMITH, Rachel, and Tamara Rathcke. 2020. Dialectal phonology constrains the phonetics of prominence. *Journal of Phonetics* 78: 1–17.

STANFORD, James N. 2008. A sociotonetic analysis of Sui dialect contrast. *Language Variation and Change* 20: 409–450.

SUN, Jackson T.-S. 1997. The typology of tone in Tibetan. In *Proceedings of the Symposium "Chinese Languages and Linguistics IV: Typological studies of languages in China"*, 485–521. Taipei: Taiwan.

SUN, Jackson T.-S. 2003. Variegated tonal developments in Tibetan. In *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, edited by David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood, 35–51. Canberra: Pacific Linguistics.

SUN, Jackson T.-S. 孙天心. 2005. Jiarongyuzu yuyan de yin'gao: liangge ge'an yanjiu 嘉戎语组语言的音高：两个个案研究 (On Pitch in the rGyalrongic Languages: Two Case Studies). *Yuyan yanjiu* 语言研究 (Language Research) 25(1): 50–59.

SUN, Jackson T.-S. 2008. Tonality in Caodeng rGyalrong. In *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag (Chomolangma, Demawend and Kasbek: Festschrift for Roland Bielmeier on his 65th birthday)*, edited by Brigitte Huber, Marianne Volkart, Paul Widmer, and Peter Schwieger, vol. I, 257–280. Halle (Saale): International Institute for Tibetan and Buddhist Studies.

SUOMI, Kari, Juhani Toivanen, and Riikka Ylitalo. 2001. On distinguishing stress and accent in Finnish. *Lund University Department of Linguistics. WorkingPapers* 49, 152–155.

SUOMI, Kari, Juhani Toivanen, and Riikka Ylitalo. 2003. Duration and tonal correlates of accent in Finnish. *Journal of Phonetics* 31(1): 113–138.

TILSEN, Sam. 2016. A shared control parameter for F0 and intensity. *Proceedings of Speech Prosody* 2016: 1066–1070.

THOMASON, Sarah G., and Terrence Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

VAN HEUVEN, Vincent. 2018. Acoustic correlates and perceptual cues of word and sentence stress: Towards a cross-linguistic perspective. In *The Study of Word Stress and Accent: Theories, Methods and Data*, edited by Rob Goedemans, Jeffrey Heinz, and Harry van der Hulst, 15–59. Cambridge: Cambridge University Press.

YIP, Moira. 2002. *Tone*. Cambdridge: Cambridge University Press.

YU, Dominic. 2012. Proto-Ersuic. PhD diss., University of California, Berkeley.

WANG Zhijie 王志洁，and Feng Shengli 冯胜利. 2006. Shengdiao duibifa yu Beijinghua shuangyinzu de zhongyin leixing 声调对比法与北京话双音组的重音类型 (Tone contrast and disyllabic stress patterns in Beijing Mandarin). *Yuyan Kexue* 语言科学 (Linguistic sciences) 5: 3–22.

WU, K. and D.G. Childers. 1991. Gender recognition from speech. Part I: Coarse analysis. *Journal of the Acoustical Society of America* 90: 1828–1840.

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

卡拉乡里汝语（藏缅语）韵律突显的声学相关物

齐卡佳
法国国家科学研究中心东亚语言研究所；南开大学文学院

侯佩好
北卡罗莱纳州立大学

Rada Chirkova
北卡罗莱纳州立大学

Angélique Amelot
法国国家科学研究中心语音学和音位学研究所

摘要

本研究考察里汝语中韵律突显的声学相关物。里汝语曾被认为拥有一种混合的韵律系统，结合了单音节词的声调，以及多音节词上的重音似的和声调式的突显格局，尽管缺乏相关的实证证据。本研究对木里县卡拉乡里汝语不同音高模式(HL、HH、LH)的双音节词进行了声学考察。使用音长、音强和基频的参数，确定不同类型双音节词中的突显形式，并探讨这些突显与词层面声调之间的相互关系。测量数据来自 8 位里汝语母语者(4 男 4 女)的语音实验。声学结果和统计分析表明，卡拉乡里汝语的全部三种音高模式中，第一个音节都处于突显位置，HL 类型的音高模式表现得更像重音（以音强为重音的声学相关物），HH 和 LH 类型的音高模式表现得更像声调（以词调的完整实现来体现重音）。这项研究有助于更好地理解里汝语的韵律结构，并为进一步探索其他里汝语土语提供了方法。

关键词

韵律突显　重音　声调　里汝语　藏缅语族