BETWEEN TONE AND STRESS: ACOUSTIC EVIDENCE OF LEXICAL PROMINENCE IN SIRIANO

XXX

XXX
XXX

ABSTRACT

This paper provides the first phonetic study of lexical prominence in Siriano, an Eastern-Tukanoan language spoken in Colombia and Brazil, whose prosodic features are poorly understood. The language, along with its sister languages in the language family, is reported to have contrastive high (H) and low tone (L), but there is no consensus among researchers whether this is due to stress or lexical tone. The current study provides acoustic evidence indicating that certain syllables are more prominent than others within a word and shows that stress and lexical tone are independent properties of the system. While primary stress and lexical tone do not influence each other, voiceless obstruent onsets attract stress (aka onset-sensitive stress), the only factor influencing stress placement (default stress = 1st syllable, no codas allowed, no phonemic vowel length).

Keywords: Siriano, onset-sensitive stress, lexical tone, prosodic features, Eastern-Tukanoan language family

1. INTRODUCTION

The prosodic features of the Eastern-Tukanoan language family (ET languages) are still rather poorly understood. For instance, there is a long-time debate about what the contrastive pitch variation, high (H) and low (L), in lexical items in ET languages are: lexical tones [1, 2, 3], stress [4], or pitch accent [5].

Siriano, an ET language whose speakers mainly live along the Paca river and Caño Viña in Colombia and Brazil [6], is an understudied language with only less than 500 native speakers. Our current study provides acoustic evidence for Siriano being a language with with a prosodic system that contains both stress and lexical tone.

Theoretically, natural languages are divided into two categories in terms of quantity-sensitivity [7], which means whether or not a language makes use of syllable weight to determine the location of stress. Stress assignment in quantity-sensitive languages is sensitive to the structure of the syllable rime (nucleus and coda) and is usually claimed to disregard the structure of onsets [8]. Natural languages may differ in what kind of syllables they call heavy. For example, Walangama put its primary stress on CVV syllable, rather than CV or CVC syllables. Heavy syllables in this respect are the ones that contain a long vowel. By contrast, Latin prefers to assign stress on CVV and CVC, and treats CV as the light one [9]. Quantity-insensitive languages typically have fixed stress assignment position, and stress are placed on the same position in every word, such as Czech. In Siriano, codas are not permitted, and long vowels are not phonemic, which means only V and CV are allowed syllables. According to mainstream metrical theory [7], both syllable types thus have phonologically equal syllable weight.

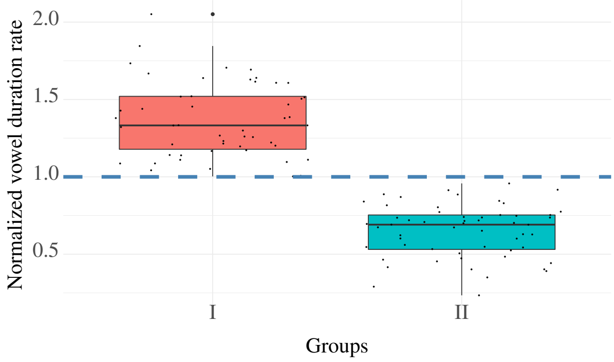

Figure 1: Relative vowel duration in disyllabic tokens with LH lexical melody.

On the other hand, we observed from elicited data that in certain disyllabic words, one syllable seems to have a longer vowel duration than the other. We divided the duration (ms) of the vowel in the first syllable by the duration of the vowel in the second syllable to normalize each word, and thus get a ratio δ. If $\delta > 1$, the first syllable has a longer vowel than the second. The result in Figure 1 shows that $\delta = 1.0$ (the blue dot line in the figure) happens to separate all disyllabic words with the same lexical melody (LH) into two groups. This indicates that in some words (GROUP I), the first vowel is phonetically longer, while other words (GROUP II) generally have a longer second vowel. This essentially rules out the case that this pitch variation is only caused by lexical tones, as lexical tones are believed not to lead to vowel duration difference [10]. Further, the onsets of the second syllable in the GROUP II words are restricted to voiceless obstruents, namely [p^h, t^h, k^h, s, h, ʔ] in Siriano. We hereby have our research questions for this study: *Does this phonetic difference in vowel duration indicate the existence of stress?* and *Does the quality of onset consonant play a role in lexical prominence in Siriano?*

In the rest of this paper, I will first introduce the methodology, including the dataset and the experiment in Section 2, then I give the results and the corresponding discussions in Section 3. Lastly, I conclude in Section 4.

2. METHODOLOGY

2.1. Speaker backgrounds and the dataset

The materials for this study are drawn from the recording of spoken Siriano that were collected as part of Wilson Silva's original fieldwork conducted in 2018. The recordings were taken as conversations between Silva and a male native speaker of Siriano. The demographic information of this speaker is unknown. The original recording consists of four parts: isolated words, phrases, sentences, and monologues. The speech time is not controlled, and the speaker's speech rate varies throughout the conversations. In order to study lexical prominence, we therefore only include the isolated word and phrase recordings, because the sentence and monologue are affected by the speaker's listing intonation. Each isolated word is repeated twice or more by the speaker. The phrases are built up with these isolated words.

2.2. Annotation and measurements

All tokens were annotated by segment by hand via Praat [11], according to the available Siriano grammar description [6, 12] and a Siriano-Spanish dictionary [13]. Since codas are not allowed, we follow the three criteria, as shown below, to determine the syllabic or lexical boundary. We look for

- a sudden change of the formant values;
- a change of the shape of the periodicity;
- a voicing offset.

Tone, either L or H, on each syllable is determined based on the F0 analysis from ProsodyPro [14], in which eight F0 values are extracted in each normalized syllabic interval; F0 sample rate is 100 Hz. The results for each acoustic parameter are subsequently subjected to statistical analysis in R [15]: a Welch's *t*-test is conducted when comparing between different groups, given that each group has different amounts of tokens. Our dataset shows that all the tokens must have at least one H tone. For all disyllabic tokens, we have 3 possible lexical melodies: LH, HH and HL, while in all trisyllabic tokens, we have 7 possible lexical melodies: only LLL is not found. The count of each lexical melody is shown in Figure 2. Since we only have plenty of data with particular lexical melodies, we put our focus on disyllabic words with LH lexical melody, and trisyllabic words with LHH, LHL and HLH lexical melodies. [1]

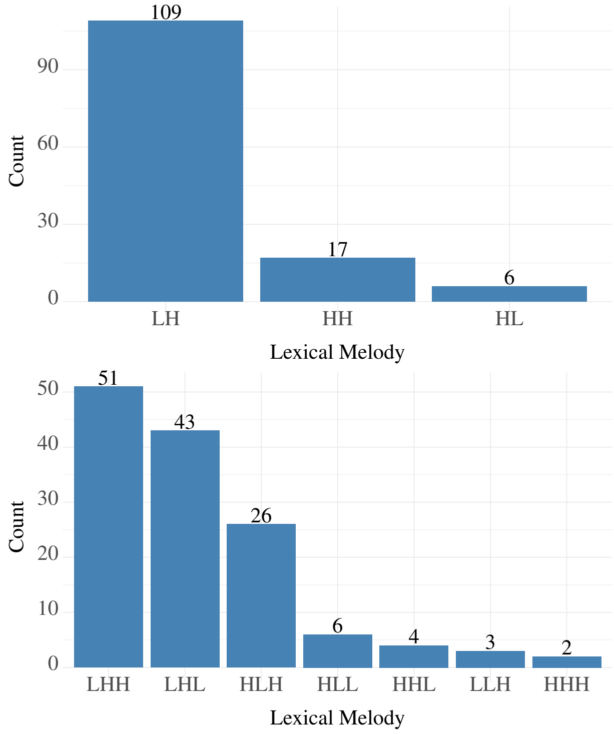

Figure 2: The count of each lexical melody based on lexical tone in the dataset for the current study. Upper: the count of the 3 possible lexical melodies of all disyllabic tokens; Lower: the count of the 7 possible lexical melodies of all trisyllabic tokens.

2.3. Procedure

All the data with certain lexical melody types are classified into groups based on the onset of the non-initial syllables. For disyllabic words with LH, GROUP II has all the words whose second syllable has voiceless obstruents as onset only, while the rest belong to GROUP I. For trisyllabic words, they have two more groups, as shown below:

- Group I: Second and third syllables do not have voiceless obstruent onset;
- Group II: Onset of the second syllable is restricted to voiceless obstruents;
- Group III: Onset of the third syllable is restricted to voiceless obstruents;
- Group IV: Onset of the second and the third syllables are restricted to voiceless obstruents.

The count of each group with particular lexical melody is given in Table 1.

	LHH	LHL	HLH
I	27	24	26
II	16	15	-
III	4	4	-
IV	4	-	-

Table 1: The count of each group with respect to each lexical melody.

3. RESULTS

F0, or fundamental frequency, is believed to be the most important acoustic correlate of lexical prominence, followed by vowel duration, vowel quality, etc. [9, 16], we have our comparisons of F0 between syllables with the same tone from different groups.

3.1. Disyllabic tokens: LH

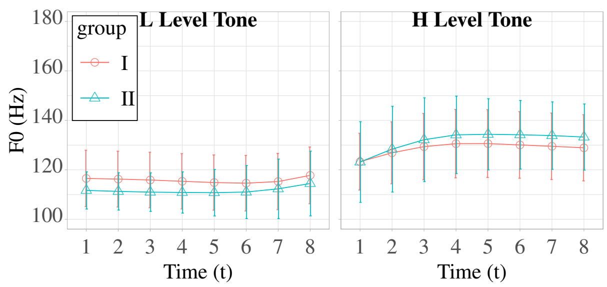

Figure 3: The F0 visualization of disyllabic tokens with LH melody.

The averaged time-normalized F0 traces (with 8 equidistant points) for disyllabic tokens with LH lexical melody are given in Figure 3. The error bars here reflect p-value of 0.05 (same below). The two panels represent the two syllables. Along with Figure 1, the first syllable (left panel) from the tokens in Group I is reliably different from that in Group II in that the first syllable has greater duration and higher F0 values on average (F0 values: $t = -7.2747$, $p < .001$), even though the first syllables of the tokens from both groups all bear L level tone. Conversely, the second syllables (right panel) in the tokens from Group II tend to have greater duration, and higher F0 values on average, even though all tokens have H level tone on the second syllable regardless of group. The comparison between the two groups in the interval, $4 \leq t \leq 8$ when values have stabilized, shows that Group II has significantly higher F0 values than Group I ($t = 10.663$, $p < .001$).

If a greater vowel duration and a higher pitch indicate metrical stress, then tokens in Group I have their primary stress on the first syllable, and stress is placed on the second syllable when voiceless obstruents are found to be onsets of the second syllable in disyllabic tokens.

3.2. Trisyllabic tokens: LHH

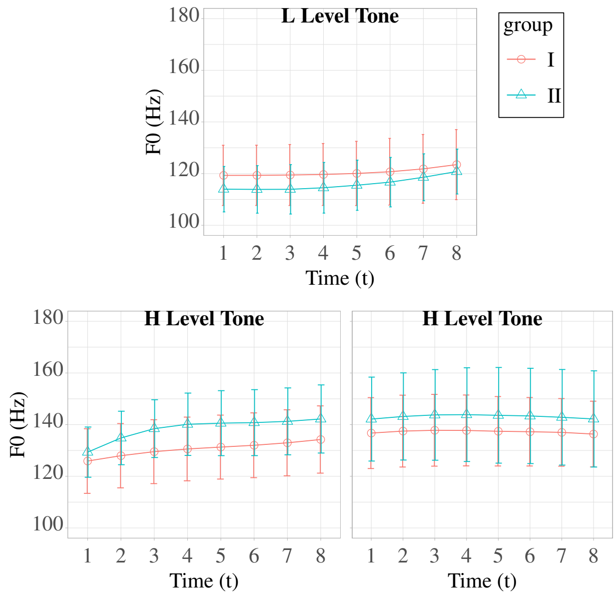

Figure 4: The F0 visualization of trisyllabic tokens with LHH melody.

Figure 4 contains average F0 traces for trisyllabic tokens with lexical melody LHH. The first syllable with level tone L is given in the top panel, while the second and third syllables with level tone H are given in the bottom ones. We can observe the same pattern here as the one we get in disyllabic tokens with LH. The first syllables bear L, under which circumstance Group I has higher mean F0 value than Group II. The reverse is also true for the second syllable. The third syllable shows the same difference as the second. In addition, in the second syllable of Group II, the F0 curve reaches the peak earlier at $t = 3$, while the F0 curve of Group I keeps climbing until the end of the second syllable, indicating that the primary stress is on the second syllable when voiceless obstruents are in the onset position.

3.3. Trisyllabic tokens: LHL

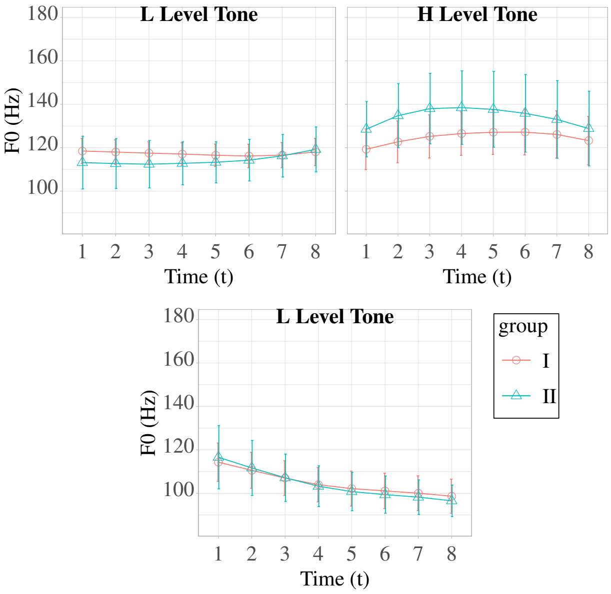

Figure 5: The F0 visualization of trisyllabic tokens with LHL melody.

Figure 5 contains average F0 traces for trisyllabic tokens with lexical melody LHL. First two syllables (top panel), with L and H tone respectively, again show the same pattern as the previous two cases. An early rise in the first syllable at $t = 6$ and a peak in the second syllable at $t = 3$ are observed as well. However, the third syllable (bottom panel) seems unexpected. The two groups (I and II) have overlapping mean F0 values ($t = 0.165$, $p = .8715$), showing no difference statistically. Compared to the first syllable with same L tone, a one-way ANOVA test shows that both curves are significantly lower than the unstressed L tone (first syllable) in GROUP II ($F = 9.01$, $p < .001$). Additional measurements, with GROUP III and IV show that when voiceless obstruents are in onset position of the third syllable, the mean F0 values of the four groups also show no significant difference ($F = 1.03$, $p = .375$), which indicates that these voiceless obstruents do not attract stress to the third syllable.

Among all 43 trisyllabic tokens with LHL, 41 tokens are bimorphemic, such as a nominal root (disyllabic) with a plural or a classifier suffix (monosyllabic). We think that in the metrical structure of a trisyllabic word with LHL, phonologically, the first two syllables from a nominal root form a foot, while the third syllable from a suffix is not a part of a prosodic word, and thus it is extrametrical, or not parsed. Consequently, when a voiceless obstruent is in the onset position of the third syllable, a stress will never be attracted out of the prosodic word.

3.4. Trisyllabic tokens: HLH

Figure 6 shows the mean F0 curves for trisyllabic words with HLH lexical melody. As shown in Table 1, tokens with HLH all belong to GROUP I, in which the two H tones, shown in left panel, are significantly different from each other ($t = -6.3938$, $p < .001$). This essentially tells us that the third syllable is primarily stressed.

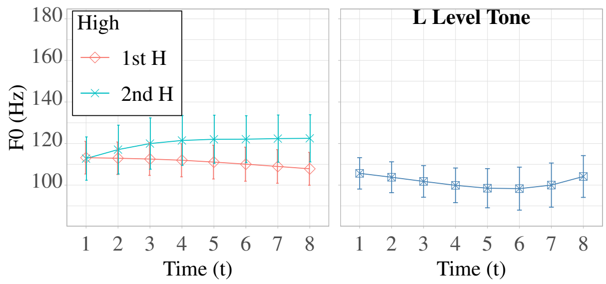

Figure 6: The F0 visualization of trisyllabic tokens with HLH melody.

4. CONCLUSION

We investigated the acoustic evidence for the existence of stress in Siriano, which is a separate system from lexical tone. Stress placement is not affected by the lexical tone type, in that both L and H tone can be stressed. The stressed lexical level tone essentially has higher pitch, compared to the unstressed one. L tone does not switch to H tone when it is stressed. Siriano is a trochaic language, with its primary stress assigned on the first syllable in a word by default except words with HLH lexical melody. A stressed syllable within a word tends to have a longer vowel duration phonetically. Our experiments show that the only effective factor influencing stress placement is the onset consonants. In particular, within a word when a voiceless obstruent in Siriano is in the onset position of the second syllable, stress is attracted to this syllable. But these consonants lose their power when they are as onsets in the third syllable of the words with LHL lexical melody. We hereby propose that Siriano is one of the languages that have onset-sensitive stress system.

5. REFERENCES

[1] E. Gomez-Imbert, “Morphologie et phonologie barasana: approche non-linéaire,” Ph.D. dissertation, Paris 8, 1997.

[2] ——, “More on the tone versus pitch accent typology: Evidence from barasana and other eastern tukanoan languages,” *Cross-Linguistic Studies of tonal phenomena: Tonogenesis, Japanese accentology, and other topics*, pp. 369–412, 2001.

[3] E. Gomez-Imbert and M. Kenstowicz, “Barasana tone and accent,” *International Journal of American Linguistics*, vol. 66, no. 4, pp. 419–463, 2000.

[4] L. M. Hyman, “Tone and/or accent. elements of tone, stress and intonation, ed. by donna jo napoli, 1-20,” 1978.

[5] J. Barnes, “Autosegments with three-way lexical contrasts in Tuyuca,” *International Journal of American Linguistics*, vol. 62, no. 1, pp. 31–58, 1996.

[6] T. Ni, “The Phonology and Morphology of Siriano: A Grammar Sketch,” Master’s thesis, The University of Arizona, 2021.

[7] B. Hayes, *Metrical stress theory: Principles and case studies*. University of Chicago Press, 1995.

[8] M. Halle and J.-R. Vergnaud, “Three dimensional phonology,” *Journal of Linguistic Research*, vol. 1, pp. 83–105, 1980.

[9] R. W. N. Goedemans, “Weightless segments - a phonetic and phonological study concerning the metrical irrelevance of syllable onsets,” Ph.D. dissertation, Leiden University, 1998.

[10] I. Lehiste, “Suprasegmentals.” 1970.

[11] P. Boersma and D. Weenink, “Praat: doing phonetics by computer (version 5.1.13),” 2009. [Online]. Available: http://www.praat.org

[12] L. Criswell and B. Brandrup, “Un bosquejo fonológico y grammatical del siriano,” in *Lenguas indígenas de Colombia: una visión descriptiva*, M. S. González de Pérez and M. L. Rodríguez de Montes, Eds. Bogotá: Instituto Caro y Cuervo, 2000, pp. 395–418.

[13] B. A. c. Brandrup, *Vocabulario siriano-español*. Lomalinda: ILV/MG, 1980.

[14] Y. Xu, “ProsodyPro—A tool for large-scale systematic prosody analysis.” Laboratoire Parole et Langage, France, 2013.

[15] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[16] D. B. Fry, “Duration and intensity as physical correlates of linguistic stress,” *The Journal of the Acoustical Society of America*, vol. 27, no. 4, pp. 765–768, 1955.

[1] The non-accessibility of the data with particular lexical melodies could be the result that, for example, LH is preferred in disyllabic words in this language.