

PHONETIC EXPONENCE OF WORD-LEVEL STRESS IN ASHANINKA (ARAWAK)

Elena Mihas^a and Olga Maxwell^b

^aUniversity of Wisconsin-Wilwaukee, USA, ^bUniversity of Melbourne, Australia

^aelmihas@uwm.edu, ^bomaxwell@unimelb.edu.au

ABSTRACT

This quantitative study examines the phonetic exponents of word-level stress in Ashaninka, an Arawak language of Peru, ISO-639-3 cni; glottocode asha1243. The variety under study is spoken in the Districts of Rio Negro, Satipo, Mazamari, and Llayla of the Satipo province, and District of Pichanaki of the Chanchamayo province of Junín Region. The analysis of Ashaninka word-level stress is based on the audio corpus of elicited speech made during focused fieldwork in the research community. The study results indicate that the right edge oriented primary stress is cued by two robust phonetic exponents, such as duration and intensity. The left edge oriented secondary stress is expressed via intensity. Vowel quality is not a statistically significant correlate to stress in the elicited data, except for the mid back /o/ whose formant structure is indicative of the levels of stress.

Keywords: Ashaninka, Arawak, word-level stress, primary stress, secondary stress, phonetic exponents.

1. INTRODUCTION

1.1. Study's significance and objectives

Ashaninka is an Arawak language of Peru, ISO-639-3 cni; glottocode asha1243. The language is a member of the Northern Kampa subgrouping of Arawak. The variety under study is spoken in the Districts of Rio Negro, Satipo, Mazamari, and Llayla of the Satipo province, and District of Pichanaki of the Chanchamayo province, Junín Region. The ethnic base of the research community is estimated to be less than 10,000 people. Most households are engaged in commercial agricultural activities. The language is spoken by the majority of adults in the parental and grandparental generations.

This research is important for two reasons. First, the study provides the first focused description of word-level stress in Ashaninka. It makes a critical distinction between word-level stress, i.e. stress in non-initial and non-final words, and phrasal stress described with reference to prosodic units higher than words. Prior publications do not acknowledge this distinction. The existing works exclusively focus on phonological analyses of segmental inventory and prosody. They include a brief description of

Ashaninka phonemes, syllable structure, and stress [5], an inconclusive qualitative study of stress [11], and an illustrative dataset of basic (morpho)phonological patterns [1]. None of these sources explores phonetic correlates of Ashaninka word-level stress.

Second, this research is the first quantitative analysis of word-level stress in an Arawak language. Across languages, stressed syllables are known to be cued by a combination of or any of the several phonetic parameters such as greater intensity (loudness), longer duration, higher pitch, and more accurate speech sound articulation in vowels vis-à-vis more peripheral vowels in stressed or secondary stressed vowels compared to unstressed vowels [6, 7, 13]. There is an overall dearth of acoustic and statistical research on the phonetic exponence of word-level stress in the indigenous languages of South America. To the best of our knowledge, there are few: on Émérillon, Tupi-Guarani, and Quechua [8, 9, 10]. The study's novel empirical data and the quantitative approach taken up to the analysis of the subject-matter render the study results highly important to South Americanists. Moreover, the study's insights are envisioned to be of significant value to prosodic typologists.

The study's narrow objective is to investigate the role of duration (§3.1), intensity (§3.2), *f0* (§3.3), and vowel quality (§3.4) in the realization of Ashaninka word-level prominence. The broad objective is to contribute to the pool of acoustic phonetic analyses of the stress systems of South American indigenous languages.

1.2. Phonemic inventory

The consonants include the voiceless stops /p, t, t̪, k/ and affricates /ts^h/ and /tʃ/, two sibilant fricatives /s/ and /ʃ/; one glottal fricative /h/; one liquid with a flap articulation /ɾ/; two glides, the bilabial approximant /β/ and the palatal glide /j/. The language has three nasal stops /m, n, ɲ/. Nasal stops /m, n, ɲ/ contrast with the placeless nasal N which occurs word-medially, following a vowel in coda position; N assimilates to a following obstruent, either a stop or an affricate. The systematic production of the alveolo-palatal stop /t̪i/ and palatal nasal /ɲ/ are observed among the speakers of the lower section of the Perené river. There are four short vowels /i, e, o, a/ and four corresponding long vowels. There are four

falling and two rising diphthongs, /ai, ao, oi, ei/ and /io, ea/ respectively. Two monophthongal vowels are front vowels, high /i/ and close-mid /e/, one low, open, near-front /a/, and one back vowel, mid-central /o/. The study investigates the phonetic correlates of the short vowels; an investigation of the long vowels will be conducted in future work.

1.3. Phonological/prosodic word (P-word)

The Ashaninka P- and M-words (P and M stand for phonological/prosodic and morphological words, respectively) are often isomorphic. Mismatches arise due to the high frequency of multiple pragmatic enclitics which behave like autonomous prosodic units forming their own stress domains. The P-word is the domain of the language-specific application of phonotactic constraints. Segmental restrictions at the P-word boundaries disallow the word-final coda and the word-initial rhotic /r/ in the onset position. At the syllable level, phonotactics restricts the minimal syllable shape V to word-initial position. A word-medial syllable could be closed by the only allowable single consonant, the placeless nasal N. The P-word minimality criterion requires two syllables to form a P-word. The P-word diagnostics include numerous phonological rules exemplified by the intervocalic voicing of stops and affricates, epenthesis of the consonantal element /t/ to break the illicit vowel hiatus, and many others. The primary stress location within a P-word is determined in terms of its orientation with respect to the word boundaries (see §1.4 for details).

1.4. Syllable structure and stress

The stress system is sensitive to the internal structure of the syllable rime. There are heavy (H) and light (L) syllables. The dividing line between heavy and light syllables falls between CVV(N), CV₁V₂(N) vs. CV(N). Heavy syllables contain branching nuclei (and optionally a coda, the placeless nasal N). Light syllables are CVN and CV.

The dominant stress patterns are penultimate and antepenultimate, with the ultimate pattern also present. The stress window is maximally trisyllabic. In disyllabic words, for the syllables of equal weight, the stress window is left-headed, (HH), (LL), e.g., /'ke.a. rio/ 'it is true', /'e.ni/ 'river'. In trisyllabic words, for light syllables the stress window is right-headed, with an extrametrical syllable present, (LL)σ, e.g., /kon.'to.na/ 'bird species', but a mixed pattern is observed for heavy syllables. In multisyllabic words, stress placement is often unpredictable: the stress window consisting of light syllables could be either left- or right-headed; heavy syllables are left-headed, H(HH), e.g., /i.ri.ma.na.ta.βa.'ka:. 'hei.tea/ 'they all will fight each other'. Words have a bidirectional rhythmic structure,

with primary stress being right edge-oriented and secondary stress located at the opposite end. Secondary stress is weight-sensitive, drawn to a heavy syllable within a trisyllabic window. When the initial and peninitial syllables are light, secondary stress tends to land on the peninitial syllable, e.g., /na.ri.ro.sa.ti.'hei/ 'we ourselves'.

2. METHODS

The current study is based on the first author's research into the Ashaninka prosodic system carried out during her multiple field trips to Peru, spanning the period of ten months. The audio corpus consists of natural discourse data and elicited word lists and sentences totaling 21 hours. The recordings were made in the villages of Impitato Cascada, Teoria, Pucharini, Rio Negro, Milagro, and Shaanki, and in the town of Satipo. The 6-hour audio recordings of elicited speech were produced by 6 literate speakers (3 females and 3 males). The elicited speech was recorded in February 2018 in the office of the local association of bilingual teachers in Satipo. All field recordings were made at a sampling rate of 44.1 kHz onto a Marantz (PMD 661) solid state recorder. The literate speakers are in the age group of 30 years and older, but none is older 50. All are bilingual in Spanish and Ashaninka. This study focuses on the examination of the elicited data; stress patterns and phonetic exponents of stress in discourse data will be investigated in the future.

The methodology included documentary fieldwork; audio recordings of the meetings with the data providers; elicitation of speakers' judgments on the word-level prominence patterns; and instrumental methods. Prior to fieldwork, lists of nouns and verbs, embedded in carrier sentences were designed. Two datasets were created: one for measuring the phonetic correlates of primary stress syllables (16 utterances), and another for examining non-primary stress in verbal words (12 utterances). The carrier sentences consisted of three words each, with the target placed sentence-medially to control for phrasal effects (phrasal prominence is attracted to left edges of phrases). All carrier sentences were neutral (non-focal). The number of recorded words and clauses averaged 160 per each person. The recorded material was transferred to the first author's laptop and sliced up into sound files using the audio editor software Audacity [2]. Acoustic analysis and annotation were carried out with the help of the speech analysis software Praat [4]. The analysis included the following measurements: target vowels duration, fundamental frequency (*f*0), intensity, and F1 and F2 taken at mid-point of the vowel in the syllables annotated as primary stressed, secondary stressed, and unstressed. The acoustic and durational characteristics of the four vowel phonemes /i, a, o, e/

were extracted with the help of a Praat script and imported into R for statistical analysis [12]. The values for the vowel tokens labelled as ‘primary stressed’ and ‘unstressed’ were extracted from the first subset (nouns), while ‘secondary stressed’ vowels were examined in verbs in clauses where verbs were non-initial. All vowels were only measured in non-phrase final syllables (in order to avoid any potential lengthening at the right boundary) and also excluded onsetless syllables. Linear mixed-effects models using the lme4 package [3] were applied to each of the measures, with random slopes of speaker and word, except for the formant measurements, where a model with speaker as an independent variable was a better fit. Likelihood ratios were used to check for validity of each model, and differences were reported with the help of Tukey’s Honest Significant Difference post-hoc test. In the case of f_0 , the data were examined separately for male and female participants to account for gender differences in the use of pitch range.

3. RESULTS

3.1. Duration

Figure 1 shows vowel duration for the three levels of stress. The likelihood ratio test shows that there is a significant effect of stress level ($\chi^2(9)=30.7$, $p<0.001$). Post-hoc Tukey confirmed that the primary stressed vowels were significantly longer than secondary stressed and unstressed vowels (S-U: $z=6.5$, $p<0.001$; S-SS: $z=3.49$, $p<0.001$), while no significant differences were found between secondary stressed and unstressed vowels. Vowel quality was also shown to have an effect: vowel /i/ was consistently shorter than vowels /e/ and /o/ in syllables with primary and secondary stress (/o-i/: $z=3.08$, $p<0.05$; /e-i/: $z=2.9$, $p<0.05$). The durational difference between unstressed and secondary stressed vowels did not reach significance; there was a large degree of inter-speaker variation in the production of secondary stressed vowels.

3.2. Intensity

Intensity patterns for the three levels of stress are shown in Figure 2. The analysis confirms a significant effect of stress level on intensity ($\chi^2(4)=30.57$, $p<0.001$). The speakers are clearly making a distinction across the levels based on this parameter. The data also consistently show that secondary stressed syllables exhibit a higher increase in intensity compared to primary stressed syllables (SS-S: $z=2.5$, $p<0.05$; SS-U: $z=5.11$, $p<0.001$; S-U: $z=5.8$, $p<0.001$). Compared with duration, this suggests that intensity is the most salient cue to secondary stress.

Figure 1: Box and whisker plot for vowel duration (ms) presented as a function of stress level (S-primary stressed, SS-secondary stressed, U-unstressed).

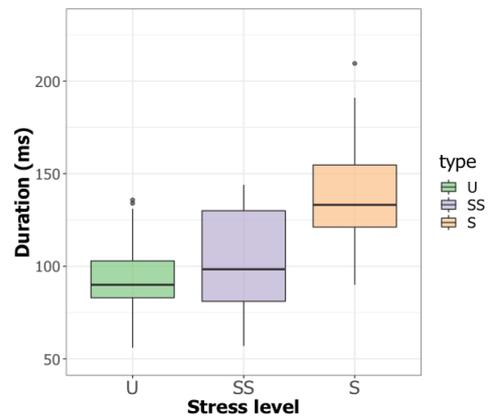
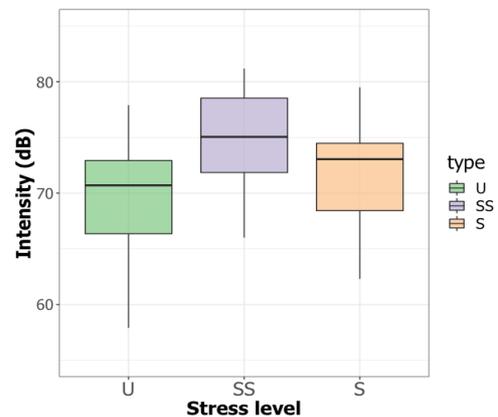


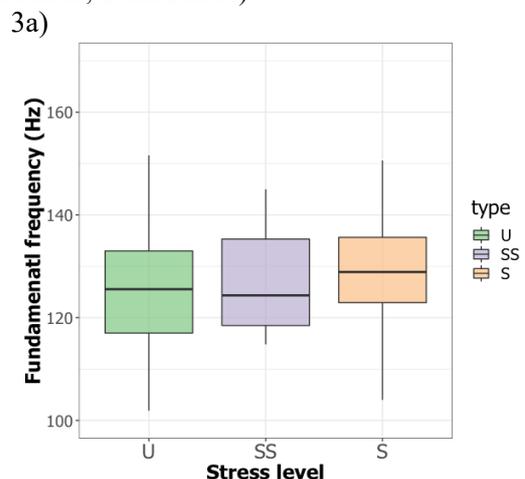
Figure 2: Box and whisker plot for intensity (dB) presented as a function of stress level (S-primary stressed, SS-secondary stressed, U-unstressed).

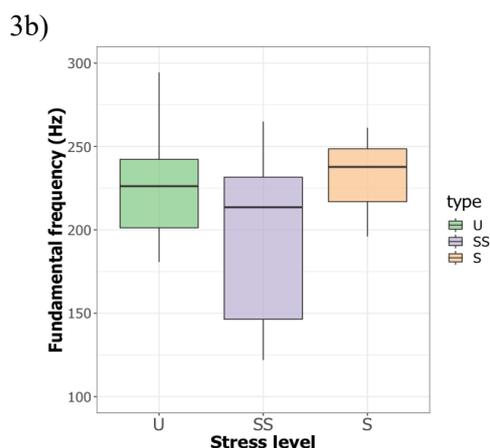


3.3. Fundamental frequency (f_0)

The results for f_0 are illustrated in Figures 3a-b.

Figures 3a-b: Box and whisker plot for f_0 presented as a function of stress level for male (3a) and female speakers (3b) (S-primary stressed, SS-secondary stressed, U-unstressed).





The emerging pattern suggests that f_0 is not a reliable cue to stress at the word level. No significant differences were observed between primary, secondary and unstressed syllables ($p > 0.05$). However, this result is inconclusive. Two male and one female speakers consistently produced primary stressed and secondary stressed syllables with higher pitch compared to unstressed syllables. A further analysis is needed that will include a wider range of tokens with secondary stressed syllables.

3.4. First and second formant frequency

The results for F1 and F2 measured at the mid-point of the target vowels are shown in Figures 4a and 4b. Compared to other acoustic measurements, F1 and F2 values demonstrate a more complex pattern of interspeaker variation. While the likelihood ratio tests found the effect of vowel and speaker for both F1 ($\chi^2(30)=92.7$, $p < 0.001$) and F2 ($\chi^2(29)=80.8$, $p < 0.001$), the effect of stress level was not significant ($p > 0.05$). However, the tests provide evidence of interaction between the stress level and the speaker, and the stress level and the vowel for both formants of the non-front vowels /a/ and especially /o/.

As can be seen in the figures, the ellipses for the front vowels /i/ and /e/ in stressed and unstressed positions exhibit an overlap in the speakers' vowel space, and thus demonstrate marginal differences in the formant structure. For the vowel /e/, there is some lowering and centralization in unstressed positions for three speakers out of six, but the differences are not robust. Several speakers produce the /i/ vowel with a higher F1 in unstressed positions, but there is no consistency. Yet in secondary stressed syllables, some speakers lower the F2 and produce a more centralized vowel /i/, as shown in Figure 4a.

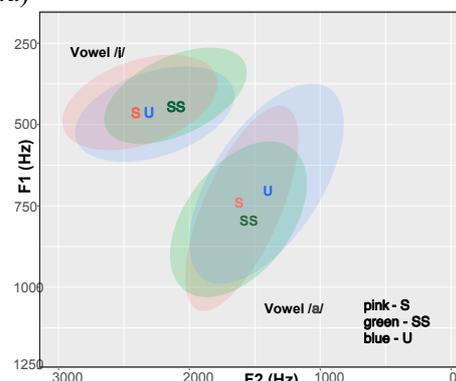
For the vowel /a/, the difference between the stressed/secondary stressed and unstressed syllables does not reach statistical significance for both F1 and F2 ($p > 0.05$). Overall, there is a great deal of intra- and interspeaker variation, yet five out of six speakers

have produced a large number of tokens with slightly lower F1 values. This result is further confirmed by the examination of formant frequencies for each speaker individually.

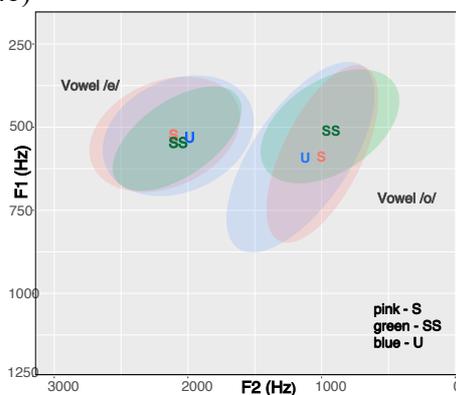
The results for the vowel /o/ are more consistent. They show differences across the three levels of stress. Higher F2 values are observed in unstressed syllables, indicating more centralized productions ($t=2.39$, $p < 0.05$), and lower F1 values in secondary stressed syllables compared to stressed/unstressed vowel productions ($t=-2.537$, $p=0.02$).

Figures 4a-b: F1 and F2 at mid-points of the four short vowels in stressed (S), secondary stressed (SS), and unstressed (U) positions.

4a)



4b)



4. CONCLUSION

The study's results have shown that Ashaninka makes clear distinctions among three levels of stress: primary and secondary stress, and lack of stress in the elicited data. Duration and intensity are robust phonetic cues of primary stress. Intensity is the correlate of secondary stress. The evidence for f_0 as one of the phonetic exponents of stress is inconclusive. The limited use of fundamental frequency as a phonetic exponent of primary and secondary stress is attested among some speakers. Formant structure is not a statistically significant correlate to stress for the non-back vowels /i, e, a/, but it has been shown to indicate the levels of stress for the back vowel /o/.

5. REFERENCES

- [1] Anderson, R. 1978. Campa Ashaninka. In: Loos, E. (ed.), *Materiales para Estudios Fonológicos*. Yarinacocha: Instituto Lingüístico de Verano, 49-82.
- [2] Audacity. Free, open source, cross-platform audio software. <https://www.audacityteam.org>.
- [3] Bates, D., Maechler, M., Bolker, B., Walker, S. C. 2015. Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1-48.
- [4] Boersma, P., Weenink, D. 1992-2017. Praat: Doing phonetics by computer (Version 6.02.24). <http://www.praat.org/>.
- [5] Dirks, S. 1953. Campa (Arawak) phonemes. *International Journal of American Linguistics* 19(4), 302-304.
- [6] Goedemans, R., Hulst, H. van der. 2013. Fixed stress locations. In: Dryer, M., Haspelmath, M. (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/14>.
- [7] Gordon, M. 2011. Stress systems. In: Goldsmith, J., Riddle, J., Yu, A. (eds.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 141-163.
- [8] Gordon, M., Rose, F. 2006. Émérillon stress: A phonetic and phonological study. *Anthropological Linguistics* 48(2), 132-168.
- [9] Hintz, D. 2006. Stress in South Conchucos Quechua: A phonetic and phonological study. *International Journal of American Linguistics* 72(4), 477-521.
- [10] O'Rourke, E. 2009. Phonetics and phonology of Cuzco Quechua declarative intonation: An instrumental analysis. *Journal of the International Phonetic Association* 39(3), 291-312.
- [11] Pike, K., Kindberg, W. 1956. A problem in multiple stresses. *Word* 12, 415-428.
- [12] R Core Team. 2015. A language and environment for statistical computing [Computer program, version 3.2.2]. Austria: The R Foundation for Statistical Computing. <http://www.r-project.org>.
- [13] Sluijter, A.M.C., Heuven, V.J. van. 1996. Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of the International Conference on Spoken Language Processing (ICSLP 96)*. www.asel.udel.edu/icslp/cdrom/icslp96.htm.