Acoustic Correlates of Word-level and Phrase-level Stress in Mankiyali

This paper investigates the acoustic correlates of word-level and phrase-level stress in Mankiyali, an endangered and highly understudied language spoken by about 500 people in the Khyber Pakhtunkhwa Province of Northwest Pakistan. Aside from the fact that a paucity of studies have been undertaken to investigate the acoustic correlates to stress in the language family and region in which Mankiyali is spoken, several additional reasons solicit the exploration of the acoustic correlates of stress in the language. First, recent surveys suggest that many of the phonological descriptions of word-level stress that we do possess may in fact unintentionally conflate descriptions of prominence at the phrase-level with prominence at the word-level (Gordon, 2011; Gordon, 2014). The present study attempts to disentangle these different levels of prominence by embedding tokens within a carrier sentence, which in turn makes up part of a three-sentence "mini-monologue", as exemplified in (1). In sentence 1, the token is new information and thus receives focus or phrasal prominence. Sentence 2 sets up the context for the token to appear in the non-focal position in sentence 3, in which the focus is on *coor* because of the pragmatic context provided by sentence 2. Thus, tokens will be non-focused and thereby devoid of phrasal stress in sentence 3, allowing for the uninhibited analysis of word-level stress.

(1) Mini Monologue frame used in the study

Sentence 1
miṇi saŋgi **[token]** *manǰu*
My friend **[token]** said
"My friend said **[token]**"

Sentence 2
miṇi saŋgi du var **[token]** *manǰu*
My friend two times **[token]** said
"My friend said **[token]** two times"

Sentence 3
miṇi saŋgi coor var **[token]** *manǰu*
My friend four times **[token]** said
"My friend said **[token]** four times"

A second reason warranting the acoustic analysis of stress in the language is that Mankiyali possesses a robust vowel inventory, all of which make phonemic distinctions in length, thereby providing fertile ground for testing the veracity of the Functional Load Hypothesis (FLH). Proponents of the FLH argue that the use of acoustic cues in other areas of the phonological domain attenuates the reliability of those same cues as accurate phonetic markers of stress (Berinstein, 1979; Hayes, 1995; Gordon and Applebaum, 2010). Since duration is the most salient acoustic correlate to stress in many of the world's languages (van Heuven and Turk, 2021), the question arises as to whether Mankiyali can call upon duration to signify stress if duration is already utilized by the language's phonology for other means. Based on the analysis conducted thus far, it seems that duration operates as an acoustic cue to stress in Mankiyali despite its use to distinguish vowel phonemes, thus contradicting the argument of the FLH.

Third, Paramore (2021) provides impressionistic evidence to demonstrate Mankiyali's utilization of the ternary stress criterion in (2). As demonstrated by the examples in (2a), stress falls on the penultimate syllable by default when all syllables in a word are equal in weight. If, however, a heavy syllable is present, stress will shift from its default position onto the heavy syllable, as shown by CVC drawing stress from penultimate CV in (2b) and CVV drawing stress from penultimate CVC(C) and CV in (2c). Nevertheless, while the examples in (2) provide sufficient evidence to verify the ternary stress criterion suggested by Paramore, he does not confirm the weight status of CVVC or CVCC syllables in his analysis, both of which occur in the language. While, he demonstrates that CVCC outweighs CV from words like "buckle" and that CVVC outweighs CVC and CVCC from words like "stubborn man" and "friendship", Paramore contends that variations in native speaker judgements and lack of sufficient data conceal the weighting relationships between both CVVC - CVV and CVCC - CVC for stress placement.

(2) Mankiyali Stress: CVVC, CVV > CVCC, CVC > CV (Paramore, 2021)

a. Default stress position

ˈa.za	“above”	*ca.ˈma.ṛi*	“skins”	*a.ni.ˈgu.gu*	“owls”
ˈgul.yoz	“grain”	*ǰan.ˈdar.yoz*	“locks”	*ˈbẽẽ.ɣii*	“roosters”

b. CVC > CV

ma.ˈčɪr	“mosquito”	*ˈbol.bo.la*	“nightingales”
ˈbaŋg.su.va	“buckle”	*so.ˈmun.da.ra*	“seas”

c. CVV > CVC(C), CV

kam.zo.ˈrii	“weakness”	*zind.ˈgii*	“life”
zid.ˈnaak	“stubborn man”	*saŋg.ˈtoob*	“friendship”

Initial quantitative analysis suggests, however, that CVVC does in fact attract stress from penultimate CVV, indicating that Mankiyali uses the more complex scale in (3) rather than the ternary scale above.

(3) Mankiyali Quinary Stress Criterion: CVVC > CVV > CVCC > CVC > CV

Recordings for the study took place in a quiet room in the summer of 2022. Participants of the study consisted of thirty native speakers of Mankiyali, most of whom live in the predominantly Mankiyali-speaking villages of Danna and Dameka, in Khyber Pakhtunkhwa, Pakistan. The experiment targeted the penultimate syllable of disyllabic words for the acoustic analysis, but the final syllable was used for the analysis of both CVCC and CVVC since finding the necessary conditions in which to analyze these two syllable types as both stressed and unstressed in the penultimate position is unlikely. These disyllabic words were grouped into near minimal pairs, one of which presumes the target syllable is stressed, the other of which presumes the target syllable is unstressed. To illustrate, in attempting to analyze the acoustic correlates of a light CV syllable, the first word takes the shape ˈCV.CV, with stress falling on the penultimate CV. To change the position of stress in the second word, a heavier second syllable, CV.ˈCVC, draws stress away from the default position. In both words, the acoustics of the penultimate CV were analyzed and compared to the other. The analysis of all five syllable types were approached in this way.

Participants read a randomized list of the tokens out loud at a normal pace. All tokens were embedded in the mini monologues discussed above and displayed in the Urdu script on a PC. In all, five pairs of stressed/unstressed tokens for each of the five syllable types were included, and participants completed the exercise twice with approximately two days between recordings. This resulted in the collection of 3,000 tokens with word-level stress and 3,000 tokens with phrase-level stress (30 speakers × 2 stress conditions × 25 target syllables × 2 stress levels × 2 repetitions per speaker). Tokens that were mispronounced are discarded. The acoustic features analyzed for both word-level and phrase-level stress in the study include duration, spectral tilt, spectral expansion, voice quality, and F0. The vowel and syllable boundary of each target syllable was labeled in Praat textgrids (Boersma and Weenink, 2015). Statistical analysis will be carried out in R (RCoreTeam, 2016) with linear mixed-effects models once the annotations are complete. Separate models will be run for each of the five acoustic properties examined in the study. The two fixed effects – presumed STRESS condition with three levels (primary, secondary, or unstressed) and SYLLABLE TYPE, which contains five levels (CV, CVC, CVV, CVCC, or CVVC) – will be measured for interaction effects: STRESS*SYLLABLE TYPE. To control for unintended influences on the main effect, random intercepts will be included in each model for speaker, word, and repetition. Thus far, tokens from six of the thirty speakers have been annotated and analyzed for vowel duration. Initial results suggest that stress status at the word-level impacts vowel duration for all five syllable types, though additional tokens from the remaining speakers are required to determine if the effects are statistical.

References

Berinstein, A. E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics*, (47).

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer (version 5.4.21)*. http://www.fon.hum.uva.nl/praat/

Gordon, M. (2011). Stress: Phonotactic and phonetic evidence. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The blackwell companion to phonology* (2nd ed., pp. 1–25).

Gordon, M. (2014). Disentangling stress and pitch accent: A typology of prominence at different prosodic levels. In H. van der Hulst (Ed.), *Word stress: Theoretical and typological issues* (pp. 83–118). Oxford University Press.

Gordon, M., & Applebaum, A. (2010). Acoustic correlates of stress in turkish karbadian. *Journal of the International Phonetic Association*, *40*(1), 35–58. https://doi.org/doi:10.1017/S0025100309990259

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Paramore, J. C. (2021). *Mankiyali phonology: Description and analysis* (Master's thesis). University of North Texas. https://digital.library.unt.edu.

RCoreTeam. (2016). R: A language and environment for statistical computing. https://www.R-project.org/

van Heuven, V. J., & Turk, A. (2021). Phonetic correlates of word and sentence stress. In C. Gussenhoven & A. Chen (Eds.), *The oxford handbook of language prosody* (pp. 150–165). Oxford University Press. https://doi.org/doi:10.1017/9781316683101.002