Phonetics-prosody interplay is implicated in typologically rare and complex morphophonology in Tohono O'odham

Daejin Kim & Robert Cruz

Department of Linguistics, University of New Mexico
daejinkim@unm.edu, rsanta10@unm.edu

ABSTRACT

This paper explores the role of focus prominence in producing nouns with morphophonology in Tohono O'odham. Plural nouns ($C_1V_1\underline{C_2V_2}C$) are formed by reduplicating the first consonant and vowel of singular nouns ($\underline{C_1V_1}C$). Each CV receives recurring strong-weak stress, signaled with *f0* and duration. We examined how speakers produce words with the potential conflict between the C_2V_2's weak stress, indicating plurality, and focus prominence, correcting lexical and grammatical information. We find that C_1V_1 has longer duration and higher *f0* than C_2V_2. With contrastive focus, speakers increased *f0* more at C_1V_1 and decreased it more at C_2V_2 while maintaining duration. Our finding implies that the language operates both stress and mora as phonological timing units (i.e., mora-sensitive stress). We suggest that speakers' meticulous control of phonetic properties indicating prosody underlies typologically complex and rare morphophonology in Tohono O'odham.

Keywords: Tohono O'odham, CV reduplication, metrical stress, intonational prominence, mora-sensitive stress

1 INTRODUCTION

Tohono O'odham [1]–[5] is a Uto-Aztecan language spoken in South-Central Arizona (near Tucson, Arizona) in the United States and Northern Sonora in Mexico. The language is actively spoken in the community (15,000 English-bilingual speakers in 2007) but has been considered endangered [6]. The language's typologically rare CV reduplication and metrical stress phonology have been widely discussed in [7]–[15]. However, there have been no perspicuous accounts explaining how the interacting morpho-phonology is linguistically structured. We instead explored the morphophonology of the language differently by examining how speakers employ intonational (focus) prominence in producing words. This allowed us to examine how the morphophonology of the language is phonetically realized.

1.1 CV Reduplication

Previous studies [12]–[14] reported that Tohono O'odham marks plural agreements of nouns, adjectives, and verbs by copying the first CV sequence of a singular noun (e.g., *ban* 'a coyote', *gogs* 'a dog'), which is the base, and pasting the copied sequence after the base, which becomes the reduplicant of a plural noun (e.g., *baban* 'coyotes', *gogogs* 'dogs'). This process is called (full) CV reduplication [16]. Note that the language never permits the first syllable with the coda to be the base; only the first CV (or C) is permitted to be the base[i]. We argue that comparing the production of singular and plural nouns provides a good basis for understanding the morphophonology of the language. This is because the base and the reduplicant are equivalent morphophonological units with major grammatical functions, but these morphemes vary due to the metrical stress of the language.

1.2 Strong-weak metrical stress

In Tohono O'odham, strong-weak (SW) metrical stress occurs across words and phrases [7]–[11]. The stress pattern of the language has been considered trochaic with a left-headed foot structure [10]–[13]. The primary strong stress is always applied to the first vowel of a word (e.g., *kóji* 'a pig', *nówiyù* 'a cow'). The second vowel receives weak stress (or gets unstressed), and then the secondary strong stress is applied at the third vowel of words (e.g., *kókojì 'pigs*', *nónowìyu* 'cows'). Word stress in this language is primarily realized with varying fundamental frequency (*f0*) and duration of segments (**Figure 1**) [17], as what other studies similarly found in various

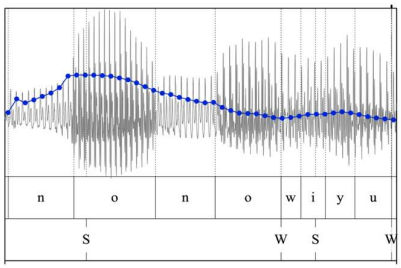

Figure 1: An *f0* contour of *nónowìyu* 'cows' with a strong-weak metrical stress in Tohono O'odham (from our recorded data).

languages. Because the language permits a CV to receive stress, the language has been claimed to have a *mora* [12], [13]. However, the metrical stress of the language has been regarded as phonologically complex because the stress pattern varies according to the status of a word, whether it is derived or not [11], and the interface with its morphology [10], [14]. In this study, we first introduce the phonetically analyzed data of Tohono O'odham, which consists of mini conversational dialogues with various words and focus prominence conditions, so we expect to gain better insight into certain aspects of Tohono O'odham's metrical stress phonology and morphology.

1.3 Intonational prominence in Tohono O'odham

Speakers actively employ prominence [18], [19] in their intonation to meet the communicative needs in conversation, such as presenting new meanings to listeners and implying speakers' intent. By comparing one element being prominent and distinctive to another, several studies cited in [18], [19] have found elements that contribute to organizing the phonetic and phonological structure of languages. Additionally, since word-level stress co-occurs with post-lexical prominence, it is necessary to distinguish stress behavior at two different levels to better understand the stress system of languages [20]. In Tohono O'odham, the realization of word-level stress should be different from that in languages previously reported because of its potential conflict with morphemes that serve a grammatical function, especially when a word needs to be contrasted lexically and grammatically in context.

As discussed in the previous section, the first CV, the base of singular and plural nouns, is subject to the assignment of strong stress. The second CV, the reduplicant of plural nouns, is subject to the assignment of weak stress. Speakers should produce the second CV prominently in their intonation when contrasting words lexically and grammatically by increasing (or decreasing) *f0* and duration, but it is under the influence of weak stress. It is unclear how language speakers produce a weakened CV sequence when it needs to be prominent due to focus.

- **Q1**: How do they produce the reduplicant (the second CV) with weak stress of plural nouns when contrasted with singular nouns?

To test this potential conflict, we compare the production of words with three types of focus prominence – broad focus (BF), narrow focus (NF), and contrastive focus (CF) [21] – which can also be implemented in Tohono O'odham. BF provides listeners with the whole phrase and sentence as new information in context, and NF narrows the range of attention to a word or a phrase as new information. CF, which has a stronger effect on articulating words and segments than BF and NF, provides a basis for correcting old information from questions (see **Table 2** for sentences designed for deriving each focus condition). In Tohono O'odham, CF can be implemented lexically (LCF) (e.g., *dog* vs. *cat*) as well as morphologically (MCF) (e.g., *dog* vs. *dogs*), which contrasts the plurality of the same lexical item. We, therefore, hypothesize that words with LCF and MCF are produced more prominently (*higher* in *f0* and *longer* in segmental duration) than those with BF and NF, and, therefore, words with MCF may have longer reduplicants in duration than those with LCF.

- **Q2**: How do speakers differentiate word prosody when lexically and morphologically contrasted in their intonation?

Taking these questions together, this study examined how speakers of Tohono O'odham produce singular and plural nouns in various focus prominence conditions, compared duration and *f0* of CVs, and explored how speakers produce a word with the potential conflict between metrical stress and focus prominence.

2 METHODS

We used eight common words (animals) as test words (**Table 1**), which consist of oral and nasal stop consonants (C = [p, t, k (fortis), b, d, g (lenis), m, n]) and vowels (V = [i, ɑ, o, u]) [1]-[4]. Test words have one to four CVs with or without a coda consonant. All plural nouns have more than two CVs. The dialogue (**Table 2**) was designed to make words be produced with various focus types. Words with LCF are contrasted with other lexical items in questions (e.g., *gogogs* vs. *totowa*) in Table 2. Words with MCF are contrasted with the same lexical items with and without CV reduplication (e.g., *gogs* vs. *gogogs*).

The dialogue was recorded in a sound-attenuated booth at the Linguistics Graduate Lab at University of New Mexico. We used a Zoom H6 recorder and a matched pair of two Rode M5 microphones. Two male native speakers of Tohono O'odham

Gloss	SG	Syll. Str.	PL	Syll. Str.
coyote	*ban*	CVC	*baban*	CV<u>CV</u>C
dog	*gogs*	CVCC	*gogogs*[ii]	CV<u>CV</u>CC
duck	*pado*	CVCV	*papado*	CV<u>CV</u>CV
turkey	*towa*	CVCV	*totowa*	CV<u>CV</u>CV
bull	*tolo*	CVCV	*totolo*	CV<u>CV</u>CV
pig	*koji*	CVCV	*kokoji*	CV<u>CV</u>CV
cat	*mistol*	CVCCVC	*mimistol*	CV<u>CV</u>CCVC
cow	*nowiyu*	CVCVCV	*nonowiyu*	CV<u>CV</u>CVCV

Table 1: Target words used for this study (SG (singular), PL (plural)). Underlined CVs are the reduplicants of plural words.

participated in this study. One speaker is Robert Cruz, a co-author and a native speaker of Tohono O'odham. During the recording, speakers were presented with conversational dialogues with questions and answers written in the language's orthography [3]–[5] and English translation on a computer monitor. Both speakers are literate in the language's orthography, and they read both questions and answers. In total, 512 word-tokens were collected from two speakers; 383 word-tokens were analyzed after excluding tokens with improper renditions.

Acoustic properties were measured using Praat [22]. Consonant duration was calculated based on closure duration and voice onset time for oral stop consonants and nasal murmur for nasal stop consonants. Vowel duration was measured based on the length of vocal fold vibration shown in the waveform and the spectrogram. *F0* values were taken from the highest values within V_1 and V_3 intervals and the lowest values within V_2 and V_4 intervals to gauge the stress pattern. We performed a series of linear mixed-effects regression analyses [23] in *R* [24] to estimate the statistical significance of independent variables (Focus Type [BF, NF, LCF, MCF], Number of CVs in words [1 – 4], and Word Type [SG and PL]) on dependent variables (CV duration, *f0*).

Prime and target sentences [Focus Type, Location]
Q: Hascu ia o'ohadag? (What is drawn here?) A: Heg añ ñeid mo g *gogogs* ha-huhu'id g mimistol. [*BF*, S] (What I see is that the *dogs* are chasing the cats.) Q: Hascu huhu'id g *mimistol*? (What chases the cats?) A: Heg añ ñeid mo *gogogs* ha-huhu'id g mimistol. [*NF*, S] (What I see is that the *dogs* are chasing the cats.) Q: No wuḍ g ***gogs*** (***totowa***) mo ha-huhu'id g mimistol? (Is the ***dog*** (are the ***turkeys***) chasing the cats?) A: Pi'a. ***Gogogs*** 'o ha-huhu'id g mimistol. [***MCF*** (***LCF***), S] (No. ***Dogs*** are chasing the cats.)
Q: Hascu ia o'ohadag? (What is drawn here?) A: Mimistol 'o ha-huhu'id g *gogogs* mo ap o'ohadag. [*BF*, O] (Cats are chasing the *dogs* is what is drawn.) Q: Hascu am ha-huhu'id g *mimistol*? (What are the *cats* chasing?) A: Mimistol 'o ha-huhu'id g *gogogs* 'o ap o'ohadag. [*NF*, O] (Cats are chasing the *dogs* is what is drawn.) Q: No wuḍ g mimistol mo (ha-)huhu'id g ***gogs*** (***totowa***)? (Are the cats chasing the ***dog*** (***turkeys***)?) [***MCF*** (***LCF***), O] A: Pi'a. Mimistol 'o ha-huhu'id g ***gogogs*** mo ap o'ohadag. (No. Cats are chasing the ***dogs*** is what is drawn.)

Table 2: Prime (Q) and target (A) sentences in Tohono O'odham and English Translation.

3 RESULTS

3.1 Focus prominence on CV duration of words

The model[iii] estimated that the recurring strong-weak stress pattern is signaled by CV duration (**Figure 2**). C_1V_1 (β (estimate) = 331, SE (standard error) = 13, p < .001) was longer than C_2V_2 (β = 210, SE = 6, p < .001). C_3V_3 (β = 226, SE = 7, p < .001) was longer than C_2V_2. C_4V_4 (β = 185, SE = 16, p < .001) was the shortest among other CV sequences (i.e., $\mathbf{C_1V_1 > C_2V_2; C_2V_2 < C_3V_3; C_3V_3 > C_4V_4}$). The model, however, found that focus prominence effects were not statistically significant for all singular and plural words (BF: β = 331, SE = 13, p > .1; NF: β = 340, SE = 6, p > .1; LCF: β = 340, SE = 7, p > .1; MCF: β = 337, SE = 7, p > .1) (i.e., **BF = NF = LCF = MCF** on *all* CV sequences)[iv]. Speakers signaled the metrical stress with varying duration of CVs, but focus prominence was *not* indicated by duration at all.

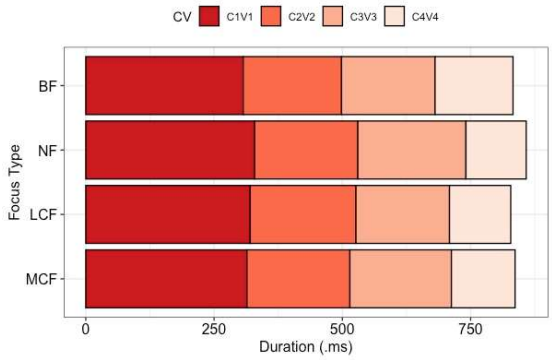

Figure 2: Mean duration of CV sequences with varying levels of focus prominence effect.

3.2 Focus prominence on *f0* of vowels

The model[v] estimated that *f0* was highest at V_1, and it gets lower towards the end of a word (V_1: β = 130, SE = 4, p < .001; V_2: β = 110, SE = 2, p < .01; V_3: β = 107, SE = 3, p < .01; V_4: β = 96, SE = 6, p < .01). Unlike the durational difference discussed in the previous sections, only V_1 and V_2 showed the statistically significant SW stress pattern. The rest of vowels did not show such a pattern (i.e., $\mathbf{V_1 > V_2 = V_3}$

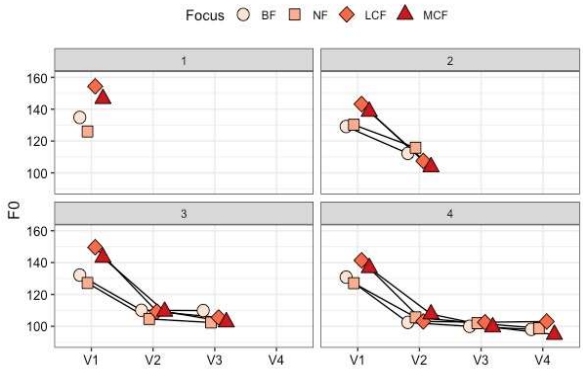

Figure 3: Mean *f0* values of vowels with varying levels of focus prominence effect. Numbers in each facet indicate the number of vowels (e.g., 1 = words with one vowel; 3 = with three vowels).

= V_4) (**Figure 3**). This may be due to the indication of a prosodic juncture signaling the end of a phrase. It can also be due to the assignment of an L tone indicating the end of a prosodic phrase [15], but we assume that the precise nature of the prosodic phrasing of the language is subject to further research.

Unlike the focus prominence on duration, the focus prominence on *f0* was realized differently across vowels. The model found that the LCF (β = 146, SE = 4, $p < .01$) and MCF (β = 139, SE = 4.5, $p < .001$) *increased f0* at V_1 as compared to BF (β = 130, SE = 4, $p < .001$) and NF (β = 128, SE = 4, $p > .01$) at V_1. LCF (β = 91, SE = 5, $p < .001$) and MCF (β = 95, SE = 5, $p < .001$) *decreased f0* at V_2 as compared to BF (β = 110, SE = 5.5, $p < .001$) and NF at V_2 (β = 112, SE = 5, $p > .1$). No contrastive focus effects (LCF and MCF) were statistically significant on V_3 and V_4. When contrasting a word, speakers enhanced word-level stress patterns by producing strong stress stronger and weak stress weaker using *f0* while maintaining its duration. Note that, as shown in **Table 3**, LCF and MCF varied depending on the number of vowels in each word. Focus prominence does not depend on grammatical status.

Word-level metrical stress is further distinguished with focus prominence, which is indicated solely with *f0*. Because speakers did not indicate focus prominence on duration, we argue that focus prominence does *not* necessarily enhance or strengthen all phonetic qualities of segments. Instead, it involves speakers' precise control of phonetic properties signaling the morphophonological structure of the language.

Words		Focus Type
1 CV (SG)	V_1	**BF = NF < LCF = MCF**
2 CVs (SG & PL)	V_1	**BF = NF < LCF = MCF**
	V_2	**BF = NF > LCF = MCF**
3 CVs (PL)	V_1	**BF = NF < LCF = MCF**
	V_2	BF = NF = LCF = MCF
	V_3	BF > NF = LCF = MCF
4 CVs (PL)	V_1	**BF = NF < LCF = MCF**
	V_2	BF = NF = LCF = MCF
	V_3	BF = NF = LCF = MCF
	V_4	BF = NF < LCF = MCF

Table 3: Post-hoc comparison of *f0* values with different focus types illustrated in **Figure 3.**

4 DISCUSSION AND CONCLUSION

We first aimed to understand how the metrical stress of the language is manifested with *f0* and duration by comparing the production of words conditioned with reduplication. Speakers employed both *f0* and duration to indicate the word-level stress pattern. Also, speakers heavily relied on acoustic cues of vowels rather than those of consonants to signal metrical stress (see endnote iv about this point).

Regarding focus prominence effects, our findings show that speakers precisely decreased *f0* of the second CV with LCF and MCF while maintaining the duration of CV sequences according to the stress pattern. In other words, when they need to contrast a word in context, they indicate more precise and distinctive metrical tones with *f0* to signal word stress. This implies that Tohono O'odham certainly treats the base and the reduplicant as equivalent morphophonological units but modifies these units according to strong-weak metrical stress using phonetic cues. We propose that Tohono O'odham uses both stress and mora as phonological timing units, and the recipient of stress is a mora (i.e., *mora-sensitive stress*), instead of a syllable, which has not been reported in languages previously studied. To better understand the co-existence of mora and stress and their interplay, further research should examine the production of words permitting C reduplication with acoustic and articulatory measurements.

We expected that speakers would produce words with LCF and MCF differently; however, MCF was statistically insignificant from LCF in both *f0* and durational measures. We assume speakers might have difficulty understanding orthographic stimuli because the language is primarily spoken rather than written. Picture-based tasks [25] may better test the effect of intonational prominence on word prosody in Tohono O'odham.

In conclusion, CV can be regarded as a basic morphophonological unit that explains lexical and post-lexical prominence variations in this language. The phonetic and prosodic modification of CV contributes to the formation of a certain aspect of morphophonology in Tohono O'odham. We look forward to further developing this study by involving more native speakers of the language; therefore, it will contribute to language documentation and revitalization of the Tohono O'odham language.

5 ACKNOWLEDGMENTS

We are highly grateful to the research participant (a native speaker of Tohono O'odham) for providing his voice and sharing in-depth knowledge of the language with us. We also thank Caroline L. Smith for making this study possible and helping us develop the study. The study is administratively reviewed and approved by the Institutional Review Board at University of New Mexico (Protocol #: 2250030063).

6 REFERENCES

[1] K. Hale, "Some Preliminary Observations on Papago Morphophonemics," *Int. J. Am. Linguist.*, vol. 31, no. 4, pp. 295–305, 1965.

[2] A. Alvarez and K. Hale, "Toward a Manual of Papago Grammar: Some Phonological Terms," *Int. J. Am. Linguist.*, vol. 36, no. 2, pp. 83–97, 1970.

[3] M. Mathiot, *A dictionary of Papago usage*, 2 vols. Bloomington, IN: Indiana University Press, 1973.

[4] O. Zepeda, *A Tohono O'odham Grammar*. Tucson, AZ: University of Arizona Press, 1983.

[5] D. Saxton, L. Saxton, and S. Enos, *Dictionary: Tohono O'odham / Pima to English, English to Tohono O'odham / Pima*. Tucson, AZ: University of Arizona Press, 1998.

[6] D. M. Eberhard, G. F. Simons, and Garcia, R. C., *Ethnologue: Languages of the world*. Dallas: SIL International, 2019.

[7] C. M. Fitzgerald, "Loanwords and Stress in Tohono O'odham," *Anthropol. Linguist.*, vol. 41, no. 2, pp. 193–208, 1999.

[8] C. M. Fitzgerald, "The morpheme-to-stress principle in Tohono O'odham," *Linguistics*, vol. 39, no. 5, pp. 941–972, 2001.

[9] C. M. Fitzgerald, "Tohono O'odham stress in a single ranking," *Phonology*, vol. 19, no. 2, pp. 253–271, 2002.

[10] C. M. Fitzgerald, "Prosodic Inconsistency in Tohono O'odham," *Int. J. Am. Linguist.*, vol. 78, no. 4, pp. 435–463, 2012.

[11] A. C. L. Yu, "Stress assignment in Tohono O'odham," *Phonology*, vol. 17, no. 1, pp. 117–135, 2000.

[12] J. H. Hill and O. Zepeda, "Derived Words in Tohono O'odham," *Int. J. Am. Linguist.*, vol. 58, no. 4, pp. 355–404, 1992.

[13] J. H. Hill and O. Zepeda, "Tohono O'odham (Papago) Plurals," *Anthropol. Linguist.*, vol. 40, no. 1, pp. 1–42, 1998.

[14] J. Riggle, "Infixing Reduplication in Pima and Its Theoretical Consequences," *Nat. Lang. Linguist. Theory*, vol. 24, no. 3, pp. 857–891, 2006.

[15] K. Hale and E. Selkirk, "Government and tonal phrasing in Papago," *Phonol. Yearb.*, vol. 4, no. 1, pp. 151–183, 1987.

[16] S. Inkelas and C. Zoll, *Reduplication: doubling in morphology*. in Cambridge studies in linguistics, no. v. 106. Cambridge: Cambridge University Press, 2005.

[17] M. Gordon and T. Roettger, "Acoustic correlates of word stress: A cross-linguistic survey," *Linguist. Vanguard*, vol. 3, no. 1, p. 20170007, 2017.

[18] F. Cangemi and S. Baumann, "Integrating phonetics and phonology in the study of linguistic prominence," *J. Phon.*, vol. 81, p. 100993, 2020.

[19] M. Grice and F. Kügler, "Prosodic Prominence – A Cross-Linguistic Perspective," *Lang. Speech*, vol. 64, no. 2, pp. 253–260, 2021.

[20] T. Roettger and M. Gordon, "Methodological issues in the study of word stress correlates," *Linguist. Vanguard*, vol. 3, no. 1, p. 20170006. 2017.

[21] C. Gussenhoven, "Types of Focus in English," in *Topic and Focus*, C. Lee, M. Gordon, and D. Büring, Eds., in Studies in Linguistics and Philosophy, vol. 82. Dordrecht: Springer Netherlands, 2007, pp. 83–100.

[22] P. Boersma and D. Weenink, "Praat: doing phonetics by computer." 2022. [Online]. Available: https://www.praat.org

[23] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using **lme4**," *J. Stat. Softw.*, vol. 67, no. 1, 2015.

[24] R Core Team, "R: A language and environment for statistical computing." Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[25] S.-A. Jun and J. Fletcher, "Methodology of studying intonation: from data collection to data analysis," in *Prosodic Typology II: The Phonology of Intonation and Phrasing*, S.-A. Jun, Ed., Oxford: Oxford University Press, 2014, pp. 493–519.

[i] The language also permits (partial) C reduplication, which copies the first consonant and pastes it after the base (e.g., *koton* 'a shirt', *kokton* 'shirts'). Some dialects permit the partial C reduplication to words differently; thus, the reduplication of Tohono O'odham varies across speakers.

[ii] The language allows complex onset and coda in the syllable [5]. [11] argued that no secondary stress is applied to odd-numbered syllables of polysyllabic words (e.g., *gogogs*, *mimistol*). Our data (or our knowledge of Tohono O'odham) indicates that the secondary stress is still applied to the odd-numbered syllables regardless of shape. Also, we see that strong-weak stress is applied to every CV of a word regardless of the derivational status of a word. Some test words are derived from Spanish (e.g., *towa* 'a turkey', *pado* 'a duck').

[iii] Duration ~ Focus Type * CV Numbers + Word Type + (1 | Speaker) + (1 | Word)

[iv] We did not include more detailed phonetic measures of CV sequences in this paper. This is because our models did not find any statistical significance of focus prominence effects on other durational measures, such as closure duration and voice onset time of consonants, and duration of vowels.

[v] F0 ~ Focus Type * CV Numbers (V) + Focus Type * Word Type + (1 | Speaker) + (1 | Word)