Acoustic phonetic properties of p-words and g-words in Sora

LUKE HORO, *Living Tongues Institute for Endangered Languages*
GREGORY.D.S.ANDERSON, *Living Tongues Institute for Endangered Languages*
AMAN SINGHA, *Living Tongues Institute for Endangered Languages*
RIA BORAH SONOWAL, *Living Tongues Institute for Endangered Languages*
OPINO GOMANGO, *Living Tongues Institute for Endangered Languages*

ABSTRACT

This paper presents the first analysis of the intonational properties of polysyllabic verbal forms in Sora, a mildly polysynthetic language belonging to the Munda language family of India. The data indicate the previous claims on Sora cannot be maintained, and the language in no sense reflects a Quantity Sensitive trochaic pattern of prominence assignment.

1 Introduction

Sora is a Munda language spoken by ca. 400,000 people mainly in southern Odisha and adjacent parts of northern Andhra Pradesh as well as in diaspora primarily in tea gardens in Assam and West Bengal, as well as Tripura, Bangladesh, etc. Morphologically, Sora can be described as mildly polysynthetic and agglutinative (in terms of grammatical words), which often consist of multiple morphemes. However, it is well known that mismatches in p-words and g-words are common in polysynthetic languages. Until very recently, g-words in Sora have not been analyzed phono-prosodically; this process is currently underway. This study is the first step to understanding the phonetic properties of g-words in Sora.

Previous work on Sora has presented a confusing array of perspectives. The original comments on the system of prominence in Sora comes from Gidugu (1931), who suggests that the pattern is not phonologically driven *per se*, but rather there is morphologically conditioned stress (or it may be conditioned morpho-lexically). In a series of papers, Donegan & Stampe (1983, 2004) and Donegan (1993) Donegan (1993) claim that Sora rather is an example of a Quantity Sensitive trochaic stress pattern, the development of which was allegedly driven by language contact at the Proto-Munda level, and this in turn triggered the alleged shift from an iambic pattern to a trochaic one. The latter claim has been proven to be inaccurate by instrumental and statistical data (Horo, 2017; Horo et al., 2020) for disyllabic words of any structure (underived, inflected). The former view of Gidugu is currently being subjected to rigorous testing and we report here some of those preliminary results.

In the following sections we detail our study. In section 2 we address what has become to be considered the traditional view of Sora of Donegan and Stampe subsumed under a heading of a theory called 'rhythmic holism' and why this cannot be maintained. In 3 we introduce the system of vowels attested in Sora. In 4 we review our instrumental conclusions on disyllables and we present an interim summary of the facts about Sora vs. the received fiction about the language. In section 5 we introduce the concept of p-words

and g-words and detail to what degree these can be shown to overlap and where mismatches might occur. Section 6 details the data collection and analyses we undertook. Section 7 presents these findings focusing on trisyllabic and tetrasyllabic g-words in Sora. Section 8 summarizes the findings.

2   The fallacy of rhythmic holism

The theory of rhythmic holism (Donegan & Stampe, 1983, 2002, 2004; Donegan, 1993) asserts that there is a fundamental rhythmic organization of languages that dictates structural typology ranging from the realization of phonemes to the organization of morphosyntax, and that this organizational parameter is crucially sensitive to language contact and macro-areal patterns, and indeed defines and explains such patterns. Sora formed the basis of this theory but has only recently been subjected to critical review from the perspective of Sora.

The foundation of the theory is that the languages that belong to the Munda and 'Mon-Khmer' groups of Austroasiatic (the latter now abandoned as a valid taxon within Austroasiatic) are claimed to canonically instantiate and exhibit South Asian [SA] vs. Mainland Southeast Asian [MSEA] typological profiles, with the latter old in Austroasiatic and the former secondary. The claim is that this was caused by a contact-triggered resetting of the rhythm of Proto-Munda from the inherited iambic/rising pattern to a South Asian trochaic/falling pattern. This shift in rhythm allegedly triggered a process of drift that subsequently entailed a complete typological shift in the South Asian Munda languages covering everything from the nature and history of the vowel systems to prosodic features to syntactic phrase structure and use of case markers in the morphosyntax. Such bold statements have been made as "….Munda and Mon-Khmer, and other South and SE Asian languages, do not just differ in structure: they are opposite at every level of structure" (Donegan & Stampe, 2002, p. 112) or "(t)he South Asian (Munda) and South-East Asian (Mon-Khmer) branches of the Austroasiatic language family are perhaps the most divergent in the world. They are opposite in structure at every level." (Donegan & Stampe, 2004, p. 3). Their claims are summarized in Table 1.

To be sure the theory has been well received and extended to other language families that also straddle South and Southeast Asia, *e.g.*, Tibeto-Burman/Trans-Himalayan. It has only recently been assessed critically in light of the attested Munda data in broader historical or comparative light (Ring & Anderson, 2018; Anderson, 2020), where it is shown that the putative dichotomy between SA and MSEA presented by the theory of rhythmic holism glosses over some actually attested facts, and agglutinative structures exist outside of SA in Austroasiatic as well, and Khasian and Nicobaric (both considered 'Mon-Khmer' in the now outdated view of the Austroasitic family) show some SA features too. As just alluded to, the understanding of Austroasiatic linguistic history has significantly advanced since the publication of these articles (Sidwell, 2014) and Munda is no longer considered to be coordinate with the rest of the family. This new research has revealed more complexities and commonalities with Munda languages than previously appreciated and reinforce the fact that Munda is simply one of many co-equal branches of Austroasiatic.

145

The basis of the theory of rhythmic holism is that Sora - and all Munda languages - underwent a shift from iambic to trochaic word structure, at the Proto-Munda language level presumably. This claim has also only recently been reviewed from the perspective of instrumental phonetics (Horo & Sarmah, 2015; Horo, 2017; Horo et al., 2020; Ring & Anderson, 2018). The results of these instrumental and statistical studies is that none of the attested acoustic cues of prominence, viz., intensity, duration and fundamental frequency, falls on the first syllable, but rather the second one in Sora disyllables. It turns out that this is also true of many other Munda languages (Santali (Horo & Anderson, 2021), Gtaʔ, Remo, Gutob). It is clear that these languages never underwent a shift from iambic to trochaic structure and therefore, nothing else in the history of these languages could be triggered by a shift that never occurred, and more nuanced, periodized and refined accounts for the development of secondary features must be advanced than a one-time resetting of 'rhythm' at the proto-language stage.

3 Sora Vowel Data

Among the claims made previously about Sora that instrumental data has revealed is that rather than a nine-vowel phoneme inventory proposed by Donegan and Stampe, Sora has six vowels, five peripheral vowels and one central vowel. These discrepancies are presented in Table 2.

The lax or lower mid and high central vowels proposed by Donegan and Stampe are attested in Sora, but they are not phonemic. As the second syllable is prominent in disyllables, there is actually a fait bit of inter- and even intra-speaker variation in the realization of the vowels of the initial non-prominent syllable, but the specific realizations of these vary considerably across speakers. This may be the result of a sampling error, with an insufficient number of speakers recorded and without being informed by processes of statistical normalization, etc., and this may have given this false impression that there are nine contrastive vowels, not the six actually attested see Figure 1.

4 Sora word prosody in disyllables

What the instrumental and statistical data reveal is that the vowels in second syllables in Sora disyllables are more peripheral and contrastive than in first syllables where they are more centralized and overlapping in vowel space, see Figure 2. As can be seen in Figure 3, 4 and 5, the three discerned acoustic cues of prominence for Sora, namely vowel duration, vowel intensity and fundamental frequency, all converge on the second syllable in Sora disyllables. Likewise, the smoothed contour graphs show that there is a clear rise in pitch on the second syllable in Sora disyllables; see Figures 6 and 7.

Thus, while the received fiction claimed that Sora has falling or trochaic word prosody, this is clearly false in Sora disyllables. Indeed, the fact is that Sora appears to retain old prosodic word patterns within its larger morphological constructs. Therefore, change in rhythm cannot explain the expansion of agglutination. Moreover, preliminary data suggest

Table 1: Munda and Mon-Khmer Structural differences (Donegan & Stampe, 2004)

	Munda	**Mon-Khmer**
Phrase Accent	Falling (Initial)	Rising (Final)
Word Order	Variable - OV, AN, POstpositional	Rigid - VO, NA, Prepositional
Syntax	Syntactic - subj/obj agreement on verb	Analytic - no inflectional morphology
Word Canon	Trochaic	Iambic, monosyllabic
Morphology	Agglutinative, Suffixing, Polysyntactic	Fusional, Prefixing or Isolating
Timing	Isosyllabic or Isomoraic	Isoaccentual
Syllable Canon	(C)V(C)	Unaccented (C)A, accented (C)(C)V(G)(C)
Consonantism	Stable, Geminate clusters	Shifting, Tonogenetic, Non-geminate, clusters
Tone/Register	Level tone (Korku only)	Contour tones or register
Vocalism	Stable, monophthongal, harmonic	Shifting, diphthongal, reductive

Table 2: Sora Vowel Data

Vowels	Donegan & Stampe (2002)	Horo (2017)	English
i	iɖ-	id-	*scratch*
ə	əɖ-	əd-	*prop*
u	-luɖ-	-luʔ-	ear
ɑ	ɑɖ-	ad-	*drive*
ɨ ∼ ə	ɨɖ-	əʔ-	*fan*
e ∼ i	-eɖ-	-iʔ-	*thorn*
ɛ ∼ e	ɛɖ-	ed-	*roll*
ɔ ∼ o	ɔɖ-	od-	*knead*
o ∼ u	-loɖ-	-lud-	*cord*
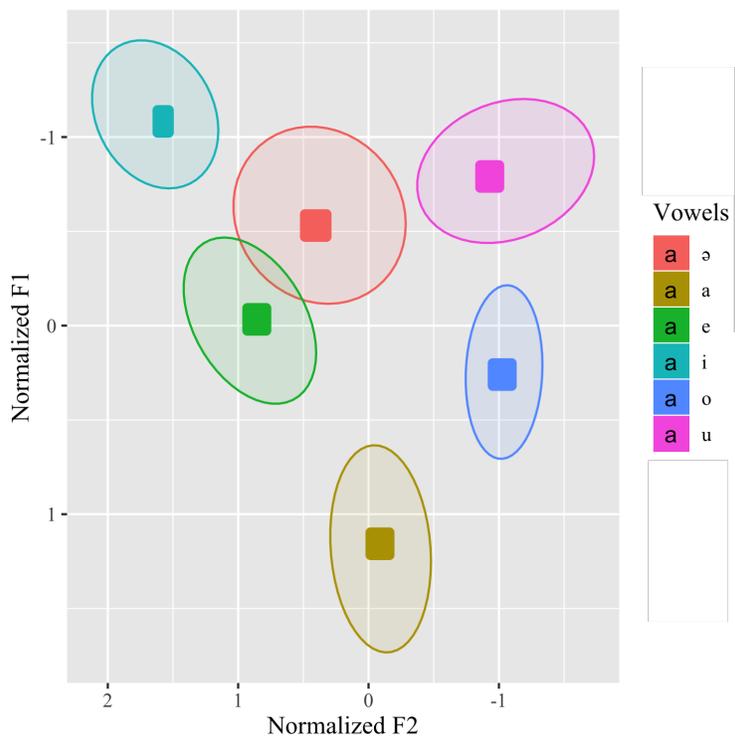
Figure 1: Sora Vowels (Horo et al., 2020)
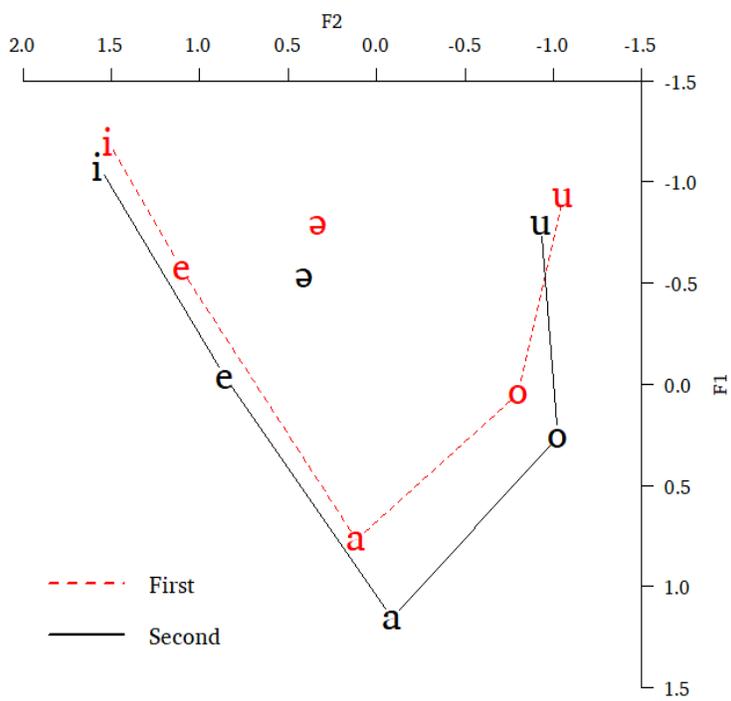
Figure 2: Sora Vowels in disyllables (Horo et al., 2020)
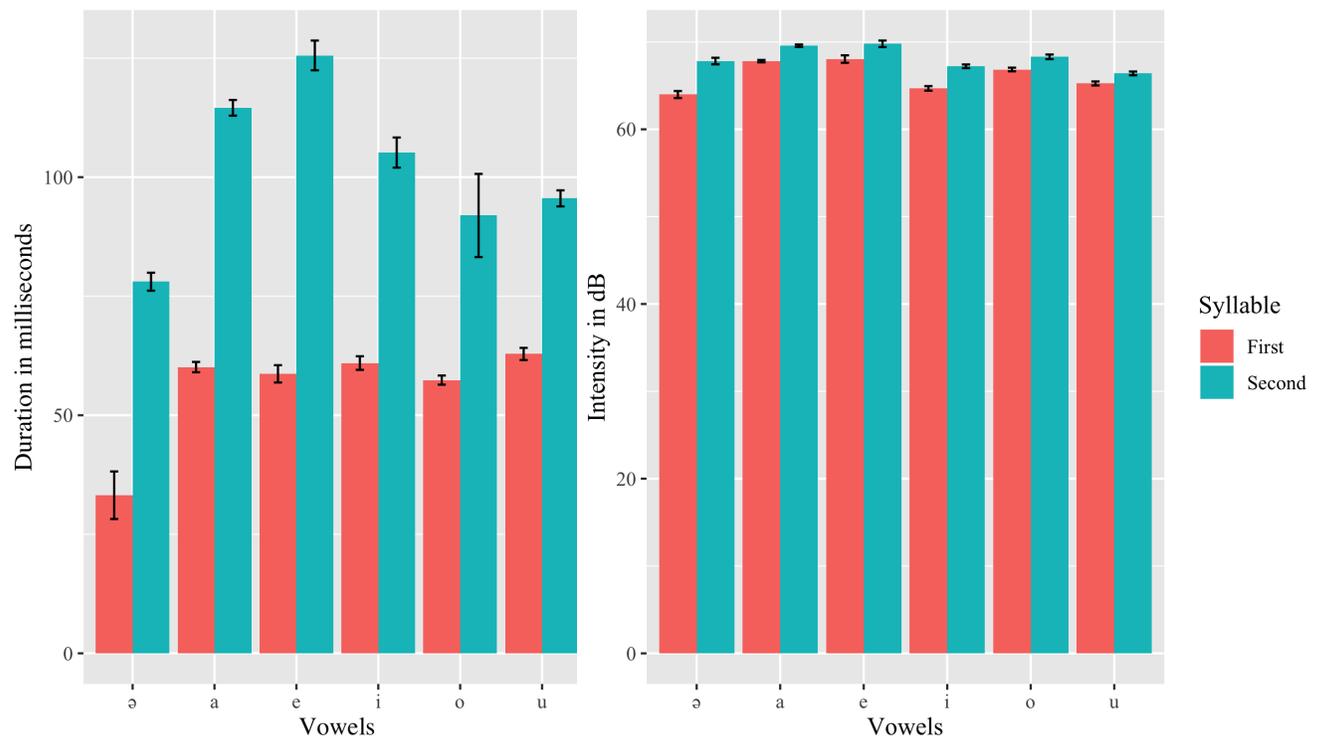
Figure 3: Vowel duration in Sora disyllablesFigure 4: Vowel intensity in Sora disyllables (Horo et al., 2020)   (Horo et al., 2020)
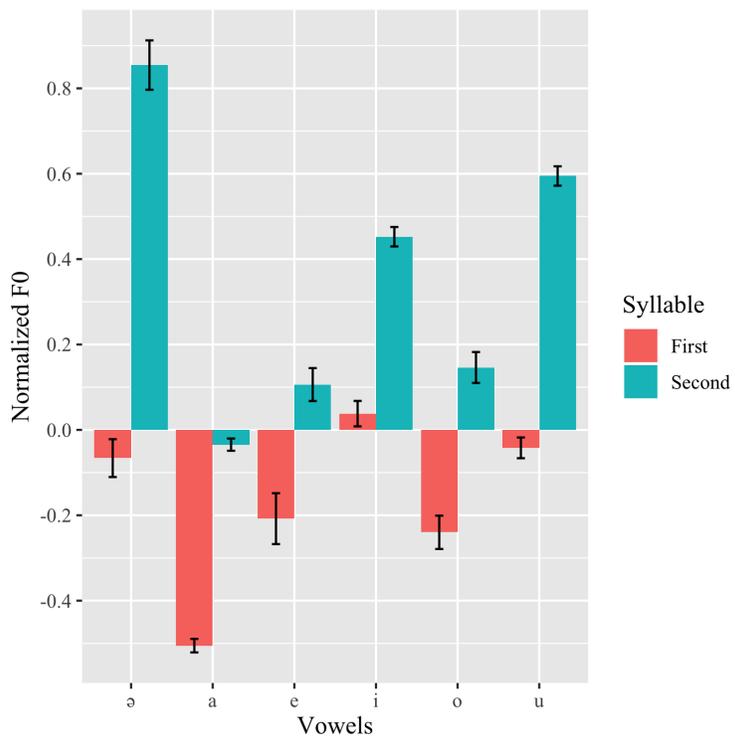
Figure 5: Fundamental frequency in Sora disyllables (Horo et al., 2020)

Figure 6: Sora pitch contour in disyllables (Ring & Anderson, 2018)

Comparison of smoothed pitch contours, Hertz:
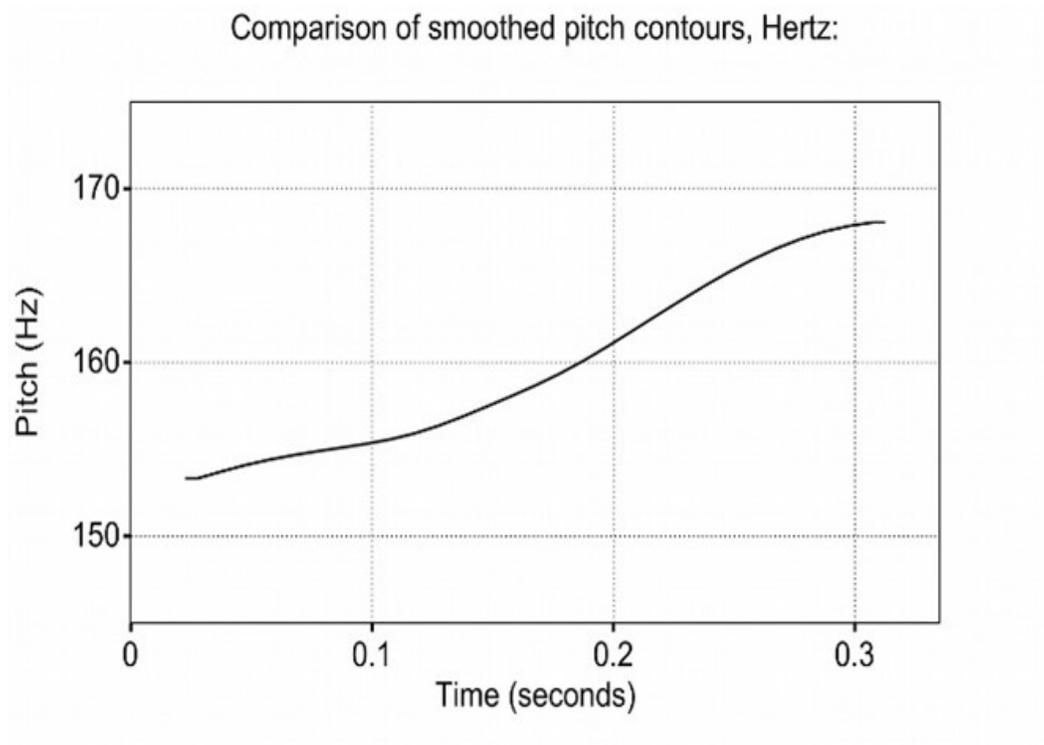
Figure 7: Sora pitch contour in disyllables (Ring & Anderson, 2018)

that many Munda languages have prominence on the second syllable, including Santali, Remo, Gtaʔ, Gutob and Gorum. The same likely is true of Koɽowa, Ho, Mundari, Korku, and Juang, thus spanning the entire genetic spectrum within the Munda family. Note that Kharia (Peterson, 2010) as well has a LH word prosody. It is of course possible that Munda languages show a low pitch first syllable prominent pattern as proposed by Rehberg (2003) for Kharia [L∗H], but the Sora acoustic/phonetic properties mentioned above favoring 2nd syllable over 1st syllable vowels do not support this.

5 P-words and G-words

It is well known that in morphologically complex languages, what is defined as a word grammatically does not always align exactly with what the prosodic or phonological characteristics of the language suggest is a word (Bogolomets & van Der Hulst, in press; Aikhenvald & Dixon, 2020; Bickel & Zúñiga, 2017; Hildebrandt, 2015). Thus, there are sometimes mismatches between p-words and g-words in such languages. One area where this is straightforward is in the placement of subject clitics in Kherwarian Munda languages such as Santali.

(1) Santali (Field Notes)

a. am iɲ=em   ɖaɽ=otʃo=ki-d=iɲ=a
    2SG 1SG=2SG.SUBJ run-CAUS=TR.PFV-TR=1OBJ=FIN
    'you made me run.'

As is seen in example 1a, in Santali, as in most Kherwarian languages, in non-imperative forms, the preferred place for the subject clitic is enclitic to the word immediately preceding the syntactic element functioning as the verb. Morphotactics may reveal the inherently ambiguous nature of some elements with respect to their phono-prosodic vs. syntactic properties in Sora. In example 2a we find what appears to be a possessive prefix, insofar as the element appears to be integrated prosodically with the following word.

(2) Sora

a. ənsəlo-n   ə-oʔon
    woman-N.SFX POSS-child
    'the woman's child.'

b. mari-n   ə-daŋgaɖi   oʔon
    Marie-N.SFX POSS-young.FEM child
    'Marie's young daughter.'

But the forms in example 2b suggest a different analysis may be warranted, or that there are phono-prosodic vs. syntactic factors at play here determining the placement of the possessive marker.

Thus, what appear to be phono-prosodic prefixes may turn out to be syntactic proclitics, even if phono-prosodically they are part of the p-word, as we see that the possessive marker

occurs on the left most part of the possessed phrase, including a modifier. These Santali and Sora data are just two subtypes of potential mismatches between phono-prosodically vs. syntactically defined 'words'.

6   Data for this study

Data for the verbal g-words of Sora used in this study were recorded in four villages of Assam: Sessa, Sinrijhan (Sonitpur District) Koilamari (Lakhimpur District) and Lamabari (Udalguri District) from forty people (ten in each village including five male and five female speakers in each location). Data were recorded for nominal g-words of Sora in 2 villages of Gajapati district Odisha (Luhangar, Luhasing) from 2 male and 2 female speakers all with no formal education and in their fifties.

The words were recorded in three contexts, i) in isolation, ii) in a phrasal frame 3a and iii) an explicitly out of focus frame 3b.

(3)   a.   ɲen _____ gamlai

'I _____ said'

b.   ɲen _____ akkarra gamlai dirga idʒdʒa

'I _____ loudly said softly did not'

Data recording was conducted in the field in a noise free environment using a Tascam linear PCM recorder and a Shure unidirectional head-worn microphone connected via XLR jack. The digital data were stored at a sampling frequency of 44.1 kHz and 32 bits in .WAV format.

7   Prominence in polysyllabic g-words in Sora

In the following sections we detail our findings on polysyllabic g-words in Sora focusing for now on trisyllabic nominal forms (7.1), tetrasyllabic nominal forms (7.2) and on trisyllabic and tetrasyllabic verb forms (7.3). Further studies currently ongoing expand this to sequences of five to nine syllables as well.

7.1   Trisyllables

Starting first with trisyllabic nouns, these are of several morphological shapes, but one potentially variable factor in compound nouns involves a combination with a shortened form of the noun used in compounds and incorporated structures known conventionally in Munda linguistics as the 'combining form' [CF]. Compound nouns can have the combining form in final or non-final position. Most are final. Other CFs are related to their corresponding syntactically freestanding and prosodically independent full forms via processes of glottal infixation or reduplication of the combining form to create the full form

[FF], not via compounding or prefixation. See Table 3 examples (i)-(iv). In the discussion that follows we refer to forms as in (iii) as ones with the CF in final position and forms like (iv) as non-final or initial.

In order to see if there was any difference intonationally in these two structural sub-types of compounds we present the data distinguishing these two subtypes. In the following graphs we also divide the data into the three different contexts of elicitation used and enumerated in 3a and 3b above. As can be seen in Figure 8, in the word in isolation context, duration peaks on the second syllable except in the quasi-focal sentential/phrasal frame in forms with the combining form in initial position. Figure 9 shows a somewhat different pattern. Here duration peaks on the final syllable in isolation, but on the second syllable in the phrasal and unaccented contexts. With respect to intensity, here we find a consistent peak on the second syllable across all three recording contexts, regardless of whether the combining form is in final position or not. See Figures 10 and 11. Fundamental frequency shows a different pattern. Fundamental frequency peaks on the second syllable on words in isolation, but on the final syllable on words in the phrasal and unaccented contexts. Like intensity this is true whether the combining form appears in final position or not; see Figures 12 and 13.

Table 3: Sora Combining Forms and Free Forms

	CF	FF	means of deriving FF from CF	gloss
(i)	-si	siʔi	glottal infixation	'hand'
(ii)	-saŋ	saŋsaŋ	reduplication	'turmeric'
(iii)	-bun	kəmbun	prefixation	'pig'
(iv)	-boŋ	boɲtel	compounding	'buffalo'
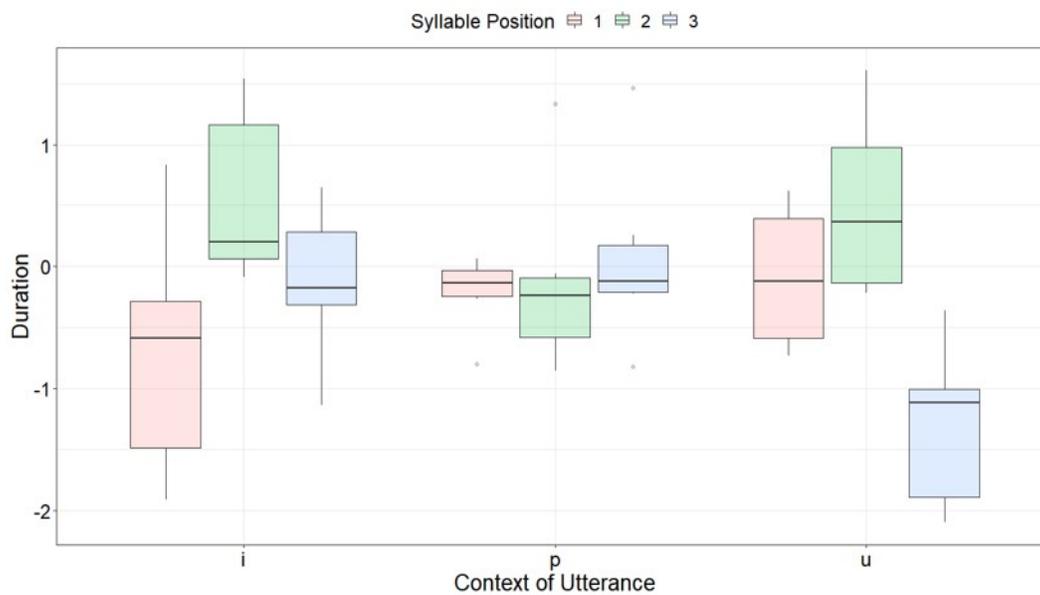
Figure 8: Vowel duration in Trisyllables
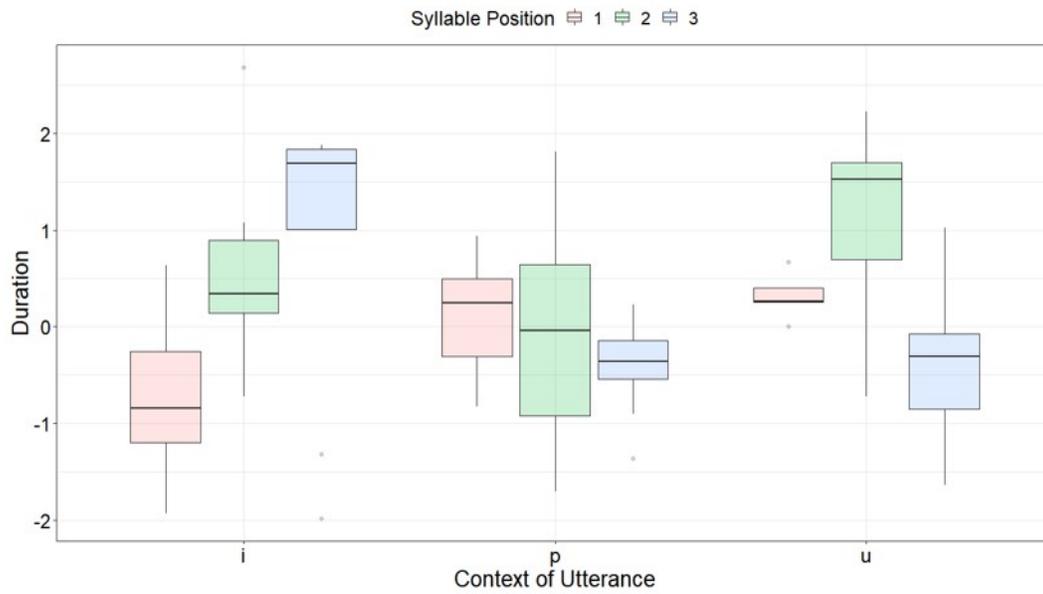
Figure 9: Vowel duration in Trisyllables
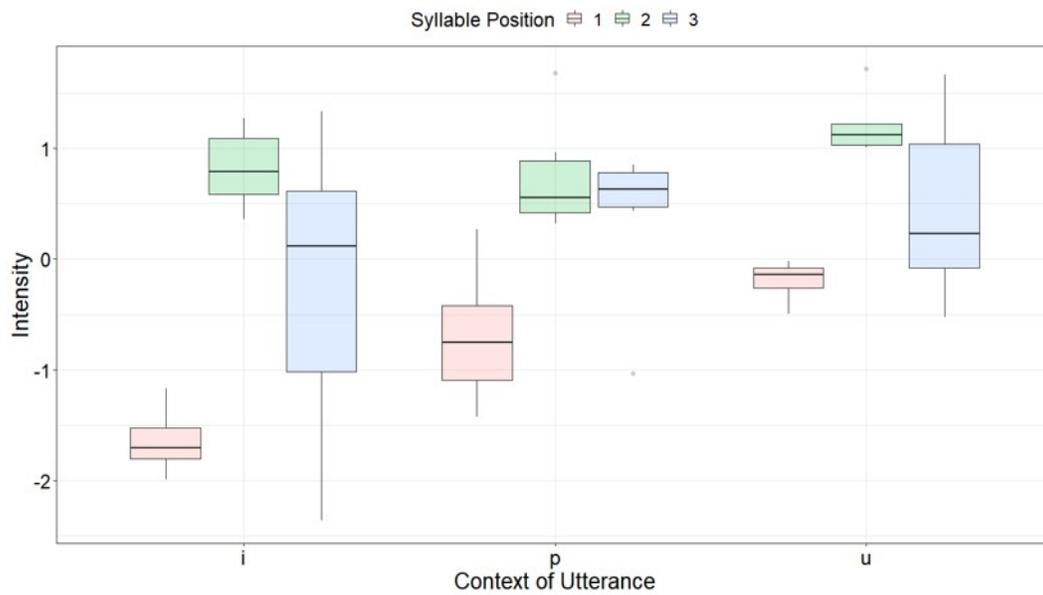
Figure 10: Vowel intensity in Trisyllables
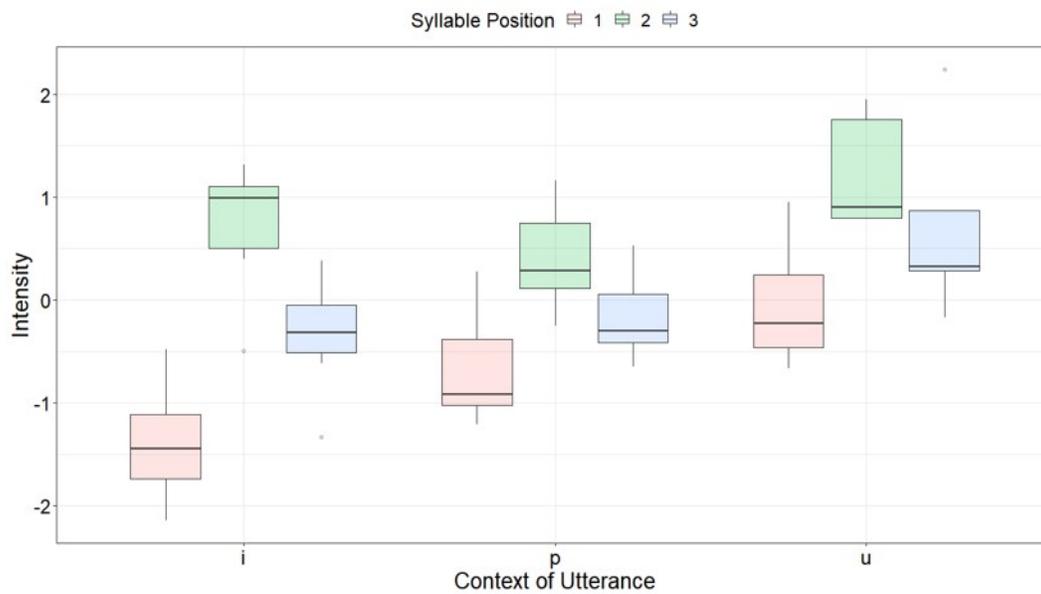
Figure 11: Vowel intensity in Trisyllables
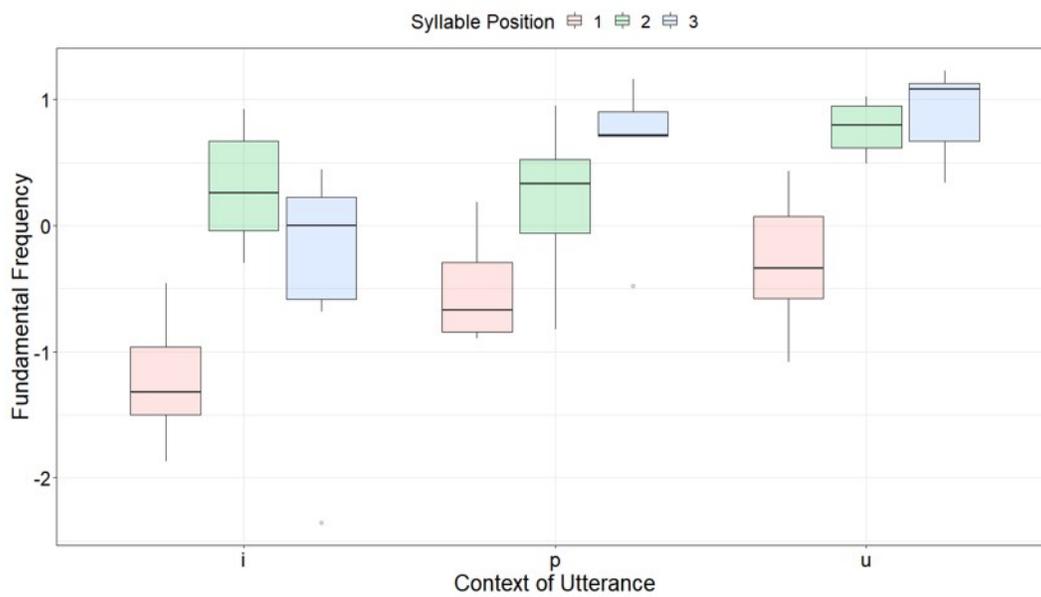
Figure 12: Fundamental Frequency in Trisyllables
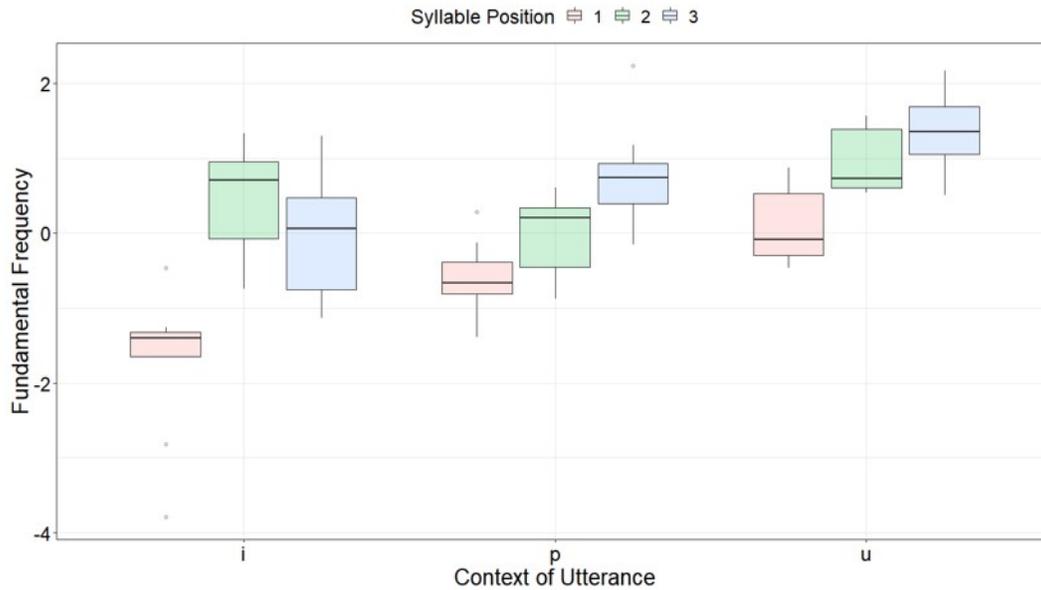
Figure 13: Fundamental Frequency in Trisyllables

7.2 Tetrasyllables

For tetrasyllabic inflected nouns, plural forms were used to keep the data consistent across all contexts. With respect to the acoustic cue of duration (Figure 14), in both isolation and the quasi-focal phrasal context we find a peak on the final syllable, while in the unaccented frame the peak is on the second syllable. Intensity patterns identically to the trisyllabic forms: Across all three contexts the peak in intensity is on the second syllable (Figure 15). Fundamental frequency in Sora tetrasyllabic nouns shows a distinct pattern. In both the isolation and unaccented contexts, the peak is on the third syllable while in the quasi-focal phrasal frame it is rather on the last/fourth syllable where the peak is typically found (Figure 16).
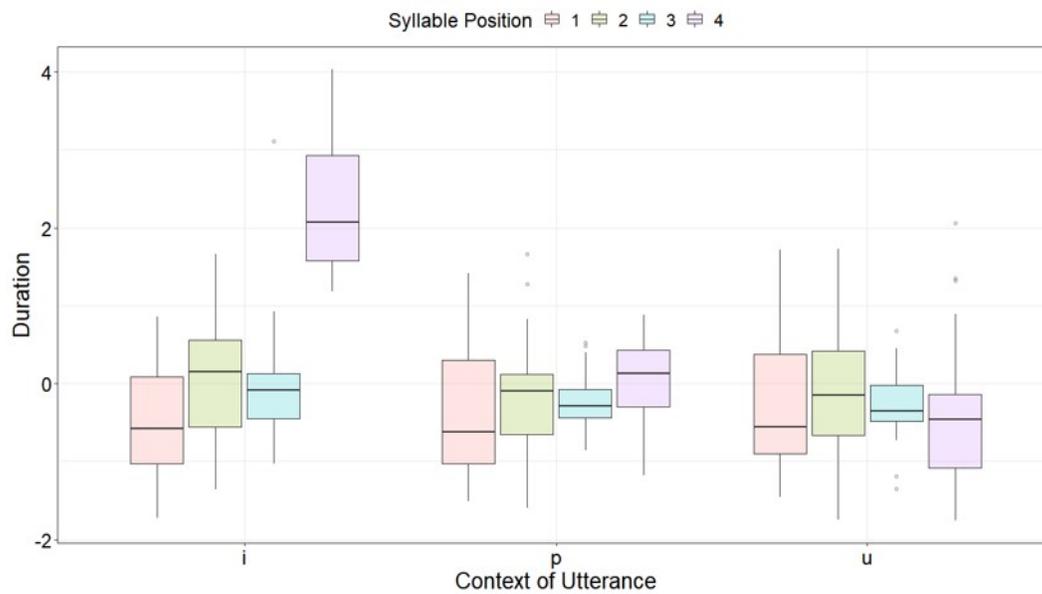
Figure 14: Vowel duration in Tetrasyllables
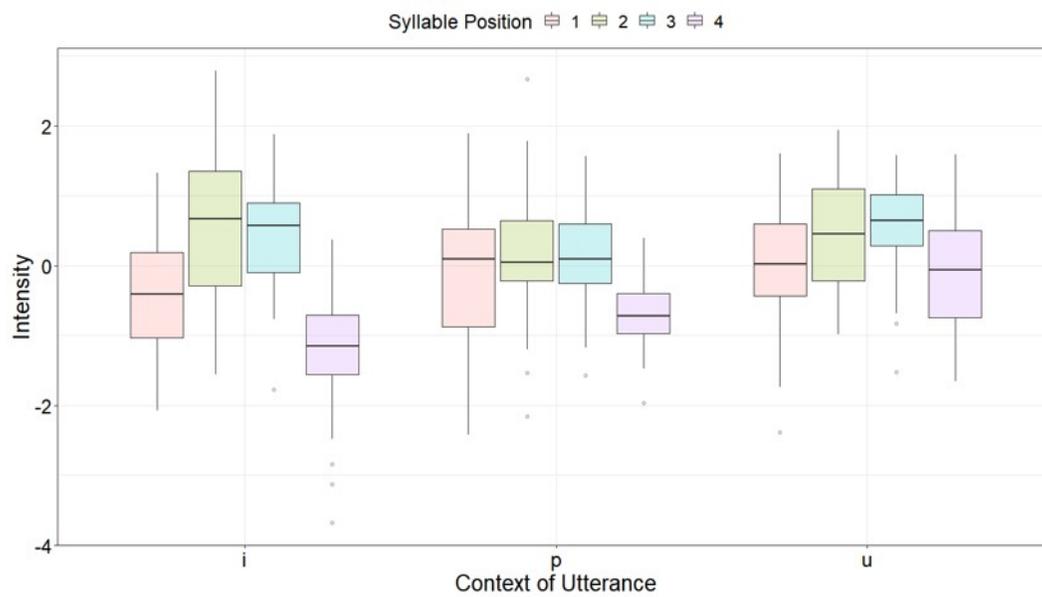
Figure 15: Vowel intensity in Tetrasyllables
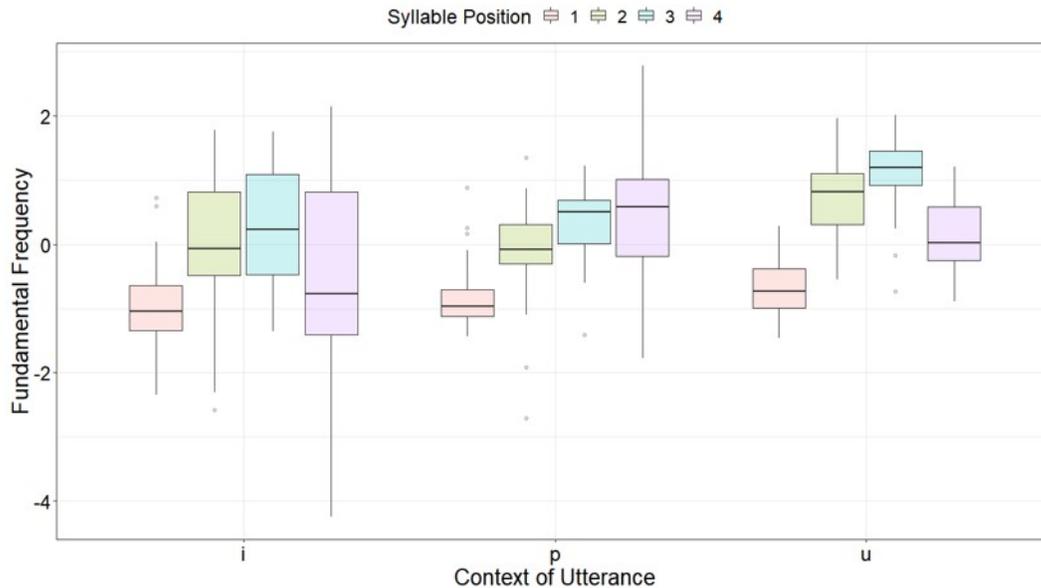
Figure 16: Fundamental frequency in Tetrasyllables

7.3 Verb forms

Before detailing the data with respect to the Sora trisyllabic and tetrasyllabic verbal forms, we first should give a very brief introduction to the Sora verb template, since as in most polysynthetic languages, it is the verb where most of the morphology is found. The verb in Sora consists of a verb stem (itself potentially simplex or derived by a voice prefix or infix), potentially proceeded by two prefix slots and up to nine suffixes or enclitic slots (see Table 4). No verb form ever has every slot filled. But given this structure, it is very easy to generate polysyllabic verb forms. For the present study we limit ourselves to discussing trisyllabic and tetrasyllabic verbs in Sora. Larger forms are presently being analyzed.

First turning to trisyllabic forms, one finds a clear peak of duration on the final syllable in Sora verbal forms of this length; see Figure 17. Perhaps unsurprisingly given the discussion of nouns above, intensity peaks on the second syllable in Sora verbs as well. This is shown in Figure 18. With respect to the pitch patterns found in Sora trisyllabic verbs, there is a clear peak in fundamental frequency on the third syllable; see Figure 19.

Turning now to tetrasyllabic verbs, we can make the following preliminary observations: The final syllable is the locus of the peak of duration in Sora tetrasyllabic verb forms; see Figure 20. Intensity shows the identical patterning to that in trisyllabic verbs and in both trisyllabic and tetrasyllabic nouns in Sora: it always peaks on the second syllable. This is graphically demonstrated in Figure 21. The pattern of fundamental frequency in Sora tetrasyllabic verb forms shows a different pattern. Figure 22 shows that the peak of fundamental frequency is on the third (penultimate) syllable. Thus, by far the most consistent cue in its distribution across these forms is the correlation of peak in intensity and the second syllable.

162

Table 4: Sora Verb Template

Position of Affixes	Function of Affixes	Position of Affixes	Function of Affixes
+2	1/2PL:SUBJ	-4	ITR/MDL
+1	NEG	-5	OBJ
0	verb.stem[1]	-6	1SG/PL:SUBJ/PST.INSTV
-1A	CF1	-7	3PL
-1B	CF2	-8	MOD/NFIN/1DL
-2	REFL	-9	COND
-3	TNS		
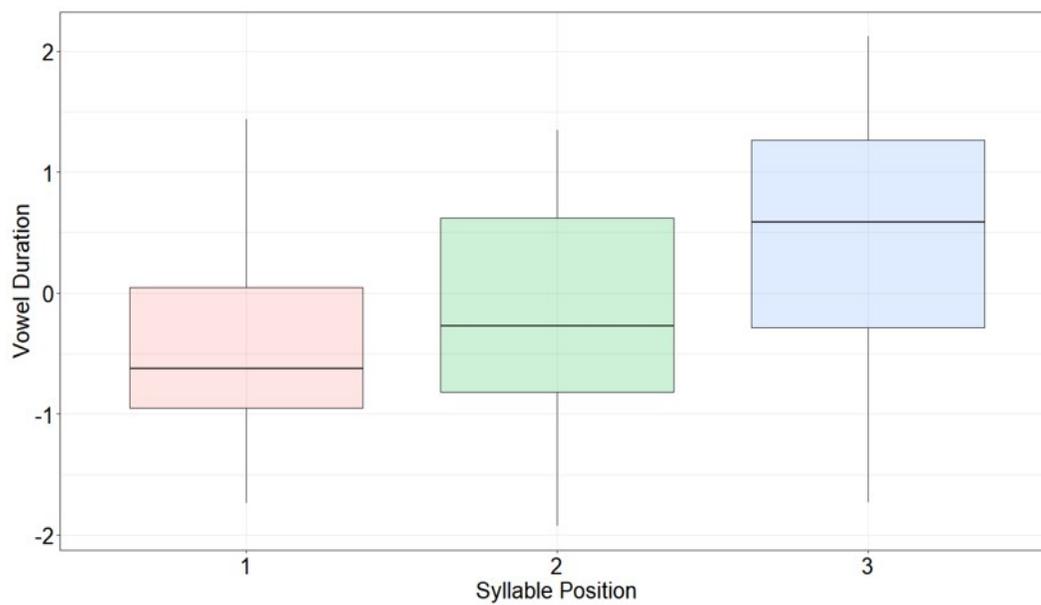
Figure 17: Vowel duration in trisyllabic verbs
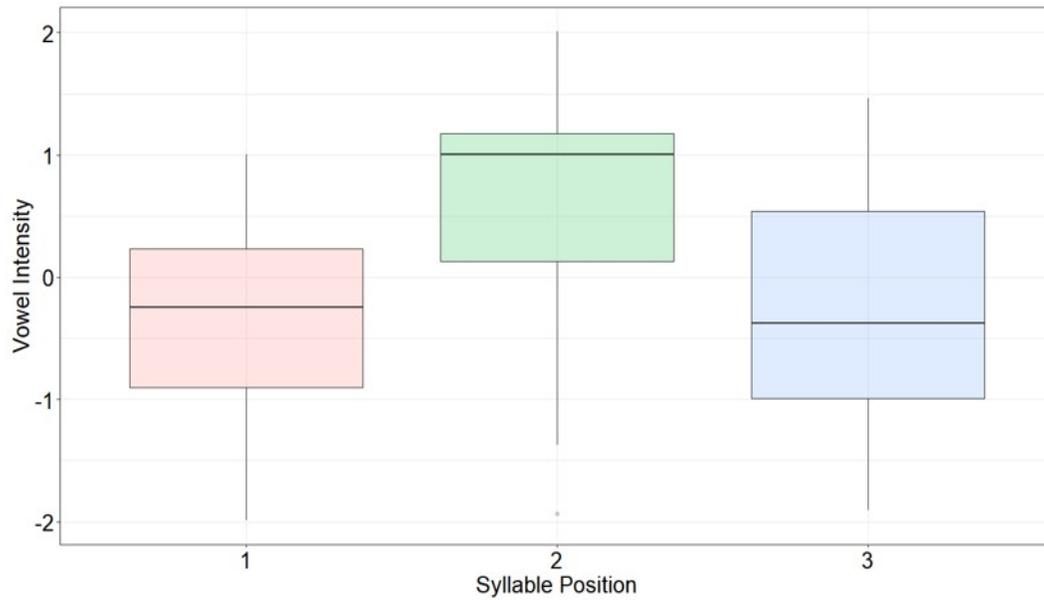
Figure 18: Vowel intensity in trisyllabic verbs
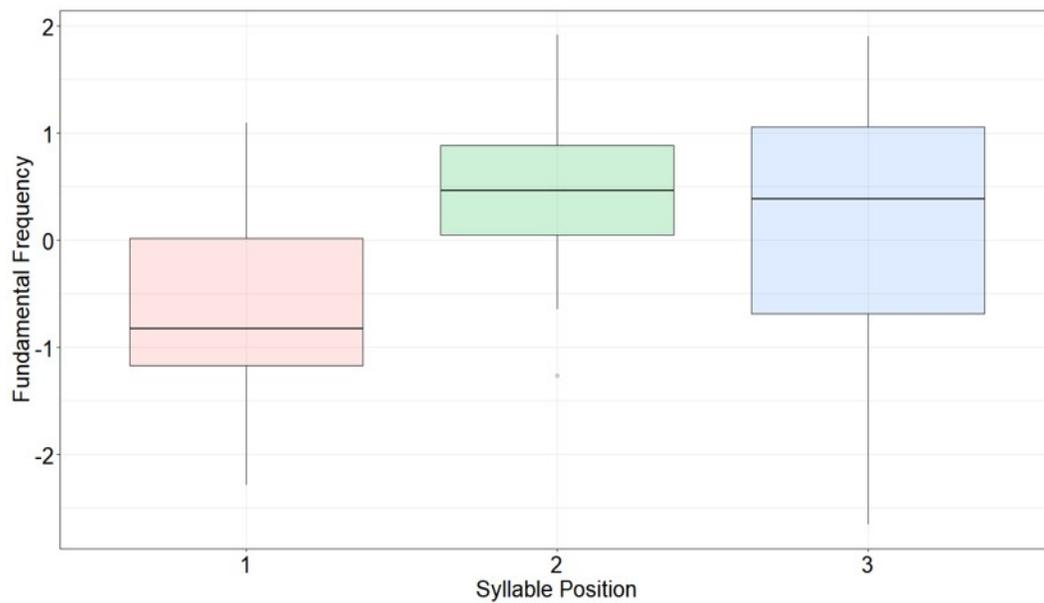
Figure 19: Fundamental Frequency in trisyllabic verbs
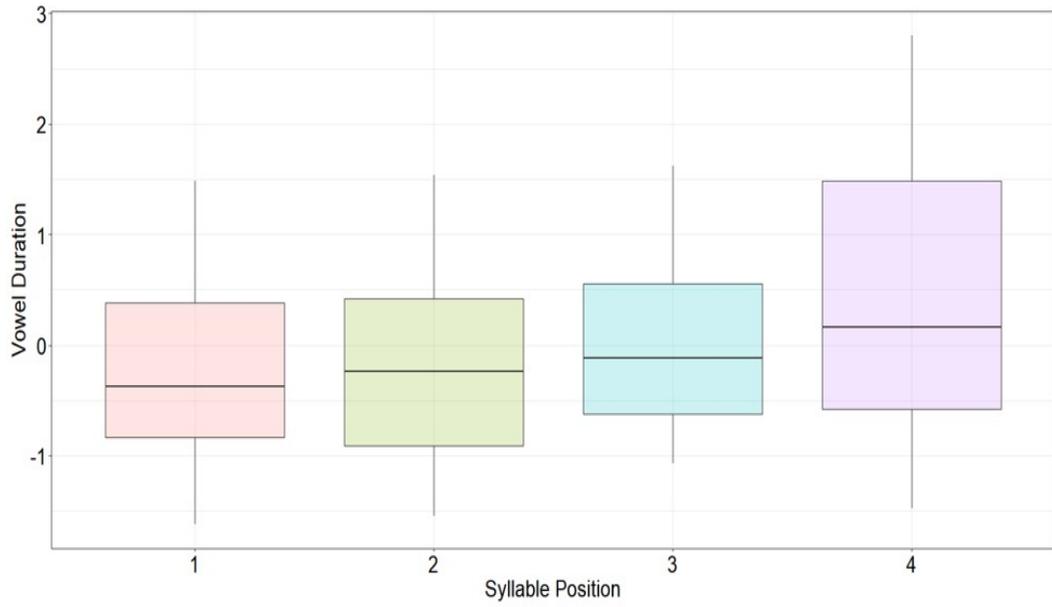
Figure 20: Vowel duration in tetrasyllabic verbs
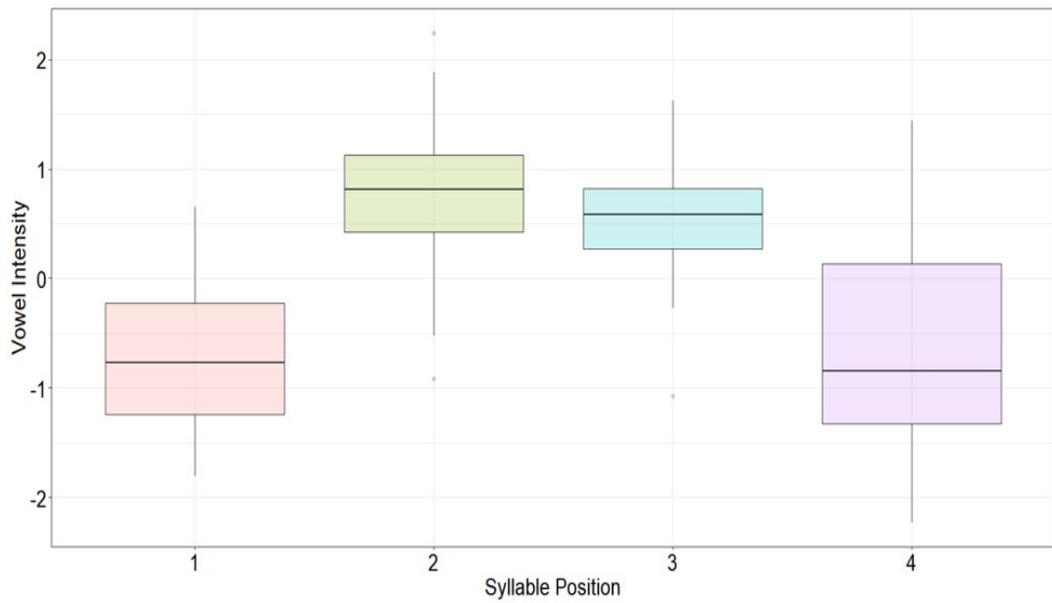
Figure 21: Vowel intensity in tetrasyllabic verbs
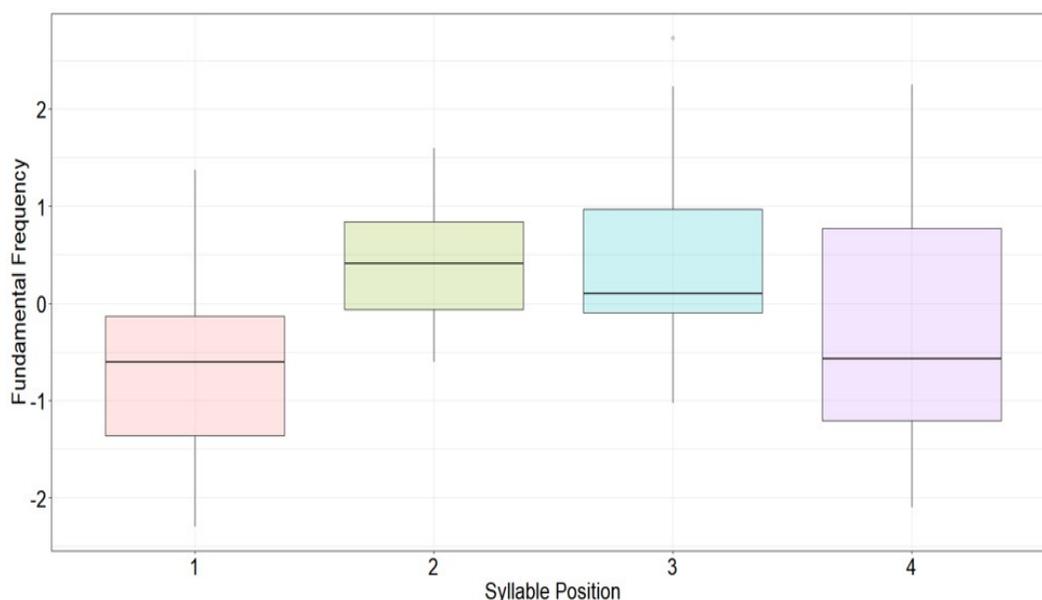
Figure 22: Vowel intensity in tetrasyllabic verbs

8 Discussion

What we can conclude from this preliminary investigation is that all disyllabic words in Sora have prominence on the second syllable seemingly cued by a conspiracy of duration, intensity and f_0. The trisyllabic and tetrasyllabic forms analysed appear to form p-words coterminous with g-words in Sora. It has also become apparent that intensity is the most consistent cue of prominence and is found on the second syllable. Duration, on the other hand, appears to delimit the last syllable, thus serving as a marker of word boundary. That fundamental frequency is found on the penult syllable in four syllable forms and may reflect a general drop in pitch in final syllable of word boundary, but it is final in three-syllable verbs, so we must still find an explanation for this, and see if larger data sets support or modify this or if this is morphologically conditioned. Verbs are morphologically complex and there may be morphemic overrides to these general patterns. A major goal of ongoing research is to resolve whether specific morphemic structure has predictable correlates to attested loci of f_0 peaks. Next steps in our research are to extend analyses to 5-8 syllable g-words to determine their intonational patterns and how these align with, or mismatch with, potential p-word patterns and to determine what is the maximal p-word in Sora.

Acknowledgements

References

Aikhenvald, Alexandra Y & R.M.W. Dixon. 2020. Words Altogether. In A. Y. Aikhenvald, R. M. W. Dixon & N. M. White (eds.), *Phonological Word and Grammatical Word: A Cross-Linguistic Typology*, 285–292. Oxford: Oxford University Press.

Anderson, Gregory D.S. 2017. Polysynthesis in Sora, with special reference to noun incorporation. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford Handbook of Polysynthesis*, 930–947. Oxford: Oxford University Press.

Anderson, Gregory D.S. 2020. Proto-Munda Prosody, Morphotactics and Morphosyntax in South Asian and Austroasiatic Contexts. In M. Jenny, P. Sidwell & M. Alves (eds.), *Austroasiatic Syntax in Areal and Diachronic Perspective*, 157–197. Leiden: Brill.

Bickel, Balthasar & Fernando Zúñiga. 2017. The 'Word' in Polysynthetic Languages: Phonological and Syntactic Challenges. In Michael Fortescue, Marianne Mithun & Nicholas Evans (eds.), *The Oxford Handbook of Polysynthesis*, 158–185. Oxford: Oxford University Press.

Bogolomets, Ksenia & Harry van Der Hulst. in press. *Word Prominence in Languages with Complex Morphologies*. Oxford: Oxford University Press.

Donegan, Patricia Jane. 1993. Rhythm and vocalic drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman area* 16(1). 1–43.

Donegan, Patricia Jane & David Stampe. 1983. Rhythm and the holistic organization of language structure. In John F. Richardson, Mitchell Lee Marks & Amy Chukerman (eds.), *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*, 337–353. Chicago: Chicago Linguistic Society.

Donegan, Patricia Jane & David Stampe. 2002. South-East Asian features in the Munda languages: Evidence for the Analytic-to-Synthetic drift of Munda. In D. Pulleyblank, J. Lawson & M. Paster (eds.), *Proceedings of the Twenty-Eighth Annual meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics*, vol. 28 2, 111–120.

Donegan, Patricia Jane & David Stampe. 2004. Rhythm and the synthetic drift of Munda. In Rajendra Singh (ed.), *The yearbook of South Asian languages and linguistics*, 3–36. Berlin, Germany: Muoton de Gruyter.

Gidugu, Venkata Ramamurti Rao Sahib. 1931. *A manual of the So:ra: (or Savara) language*. Madras, Government Press: Superintendent, Director of Public Instruction.

Hildebrandt, Kristine A. 2015. The Prosodic Word. In John R. Taylor (ed.), *The Oxford Handbook of the Word*, 221–245. Oxford: Oxford University Press.

Horo, Luke. 2017. *A Phonetic Description of Assam Sora*. Guwahati: Indian Institute of Technology Guwahati dissertation.

Horo, Luke & Gregory D.S. Anderson. 2021. Towards a Prosodic Typology of the Kherwarian Munda languages: Santali of Assam. In M. Alves & P. Sidwell (eds.), *Proceedings of the 30th Annual SEALS meeting*, 298–317. Honolulu: UH Press.

Horo, Luke & Priyankoo Sarmah. 2015. Acoustic Analysis of Vowels in Assam Sora. In

Linda Konnerth, Stephen Morey & Priyankoo Sarmah (eds.), *North East Indian Linguistics*, vol. 7, 69–86. Canberra: ANU.

Horo, Luke, Priyankoo Sarmah & Gregory D.S. Anderson. 2020. Acoustic Phonetic Study of the Sora Vowel System. *The Journal of the Acoustical Society of America* 147(4). 3000–3011.

Peterson, John. 2010. *A Grammar of Kharia: A South Munda language*. Leiden: Brill.

Rehberg, Kerstin. 2003. *Phonologie des Kharia-Prosodische Strukturen und Segmentales Inventar*. Osnabruck: Universitat Osnabruck MA thesis.

Ring, Hiram & Gregory D.S. Anderson. 2018. On prosodic structures in Austroasiatic diachrony: 'Rhythmic holism' revisited in light of preliminary acoustic studies. In Hiram Ring & Felix Rau (eds.), *Papers from the 7th International Conference on Austroasiatic Linguistics, JSEALS Special Publication 3*, vol. 11 2, 1–35.

Sidwell, Paul. 2014. Austroasiatic Classification. In M. Jenny & Paul Sidwell (eds.), *The Handbook of Austroasiatic Languages*, vol. 1, chap. 3, 144–220. Leiden: Brill.