TOWARDS A PROSODIC TYPOLOGY OF KHERWARIAN MUNDA LANGUAGES: SANTALI OF ASSAM

Luke HORO
Living Tongues Institute for Endangered Languages
luke.horo@livingtongues.org

Gregory D. S. ANDERSON
Living Tongues Institute for Endangered Languages
gdsa@livingtongues.org

Abstract
In this article, we detail the prosodic typology of disyllables in Santali spoken in Sonitpur district, Assam, India. Previous investigations present a chaotic picture of contradictory analyses. For Santali, Ghosh (2008:30) states "(s)tress is always on the second syllable of the word regardless of whether it is an open or a closed syllable". Neukom (2001:8), on the other hand, claims that in disyllabic stems, "(s)tress falls on the first syllable; however, if the first syllable is light and the second heavy (iambic structure), stress falls on the second syllable", i.e., a Quantity Sensitive system. Bodding (1922) detailed a partly morpholexical system of prominence assignment. To test this, we recorded data from male and female speakers of various ages in a noise-free environment in the field. Forms were recorded in isolation, in a quasi-focal frame "I ____ said", and in an out of focus frame "I ____ LOUDLY said, not SOFTLY" to control for focal intonation effects and those of phrasal or utterance intonation. Based on subsequent instrumental analysis, we suggest that Assam Santali always shows prominence on the second syllable of disyllables, cued by intensity, f0 and duration–a pattern like that attested in Assam and Odisha lects of Sora (Horo and Sarmah 2015, Horo 2017, Horo et al. 2020), contra Donegan and Stampe (2004)–an iambic pattern and not quantity sensitive, at least in disyllables.

Keywords: Santali, Sora, Munda, Prosody, Prominence
ISO 639-3 codes: sat

1 Introduction

In this article, we detail a preliminary study on the prosodic typology of the Kherwarian Munda languages, a group of Austroasiatic languages spoken in eastern and northeastern India, focusing here on one lect of Santali spoken in Sonitpur district, Assam, India. This variety has not been described previously, nor have any Kherwarian lects been studied experimentally using instrumental phonetic methods before.

The major group of Santali speakers resides to the southwest of this region in West Bengal, Jharkhand, and Northern Odisha. Some of these varieties of the language have been subjected to previous analysis in some domains, but not to date in terms of instrumental phonetic analysis. Thus, only impressionistic statements have been made to date in the literature about the prosodic or intonational system of Santali. Indeed, the previous investigations present a chaotic picture of contradictory analyses.

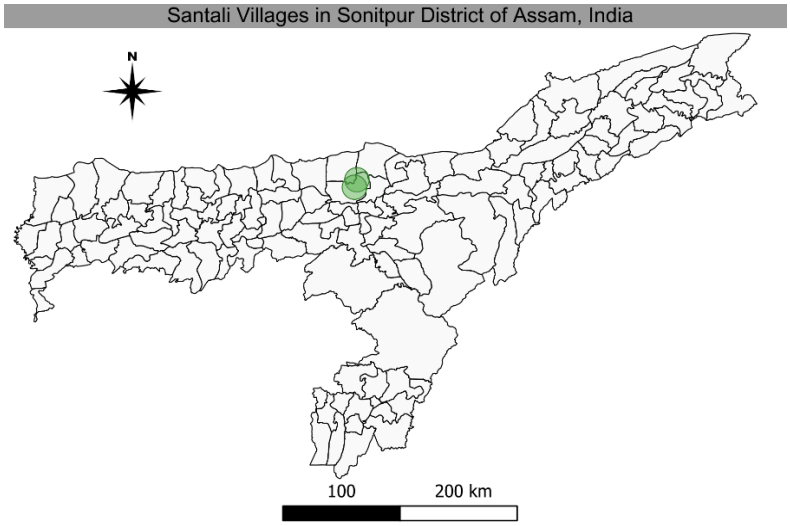
For Santali, Ghosh (2008:30) states Santali has fixed second position stress while Neukom (2001:8) claims a Quantity Sensitive system where a heavy syllable following a light one takes stress. Otherwise, it is the initial syllable. Bodding (1922) for his part detailed a partly morpholexical system of prominence assignment. Thus, subjecting this analysis to the rigors of instrumental verification is an ongoing process. In this preliminary study, we limit ourselves to uninflected lexical items that are in their underlying structure disyllabic, with a few that surface as trisyllabic with a weak medial syllable resolving a word medial onset cluster. Moreover, forms were recorded in isolation, in a quasi-focal frame "I ____ said", and in an out of focus frame "I ____ LOUDLY said, not SOFTLY" to control for

focal intonation effects and those of phrasal or utterance intonation. Based on the data acquired in the above formats and subsequent phonetic analysis, prominence in Santali disyllables is explored and the findings, supported by acoustic evidence, are reported in this paper.

2 Santali of Assam

According to the Census of India (Registrar General of India 2011) there are 7,368,192 Santali speakers in India, of which 213,139 speakers are reported in Assam. The Santali variety of Assam has not been analyzed previously. The current study presents a preliminary finding that is based on speech data recorded from Santali speakers living in two villages of Assam's Sonitpur district, namely, Erasuti and Borbil (see Figure 1). There are approximately eight hundred Santali individuals in Erasuti and one hundred and fifty Santali individuals in Borbil, and both villages have a major concentration of Santali speakers in Sonitpur District of Assam. Also, the inhabitants claim to have lived in those villages for at least four generations until now. Moreover, unlike most ethnic Munda inhabitants of Assam, who are reported to have migrated to Assam as indentured tea laborers from parts of Eastern India in the nineteenth century (Tea Districts Labour Association-India 1924), the Santali community recorded in this work claim that they are native inhabitants of the land, not migrant laborers.

Figure 1: *Map of Assam highlighting locations of Santali speech data collection*



For the purposes of this study, a total of six native Santali speakers (three male and three female) living in the two villages were interviewed and recorded to acquire the data. The average age of the participants is thirty-one years with a standard deviation of eleven years. Each participant is a multilingual speaker, and besides Santali they also speak Sadri, Assamese, and Hindi as their second, third and fourth languages. Among the six participants, only one male and one female have completed their high school education, whereas the others dropped out of their formal education either during high school or even earlier.

2.1 Data collection

To collect Santali speech samples, a dataset was created from a list of basic Santali vocabulary wherefrom, the text data, including fifty-one Santali disyllabic words having (C)V(C).(C)CV(C)[1] syllable structures, were generated. All words in the dataset are non-derived nouns, including words for

[1] Four targeted disyllabic words that have onset clusters in the second syllable, [*ʧɛpɽɛ*], [*sikɽi*], [*kʰapɽi*] and [*pandʒra*], were interchangeably produced as trisyllabic words with the insertion of an epenthetic vowel [ə] breaking the obstruent and rhotic consonant clusters present in the words. Such variants are not included in the analysis.

body parts, animal names and words for natural objects. Table 1 presents a subset of the text data used in this study to generate the speech data.

Table 1: *Subset of Santali disyllabic data used in the study*

	Santali	**English**	**Syllable Structure**
1.	*aḍo*	'urine'	V.CV
2.	*ipil*	'star'	V.CVC
3.	*supu*	'arm'	CV.CV
4.	*lutur*	'ear'	CV.CVC
5.	*ɲinda*	'night'	CVC.CV
6.	*kaṭkom*	'crab'	CVC.CVC

Subsequently, the speech data is derived from eliciting and recording the text data once in isolation and once in each of the two sentence frames shown in (1) and (2).

(1) *iŋ* ____ *men-kediŋ*
1SG ____ say-PST.TR/ACT.1SUBJ
'I said ____'

(2) *iŋ* ____ *gula=te* *men-kediŋ* *lahe=te=do* *baŋ*
1SG ____ loud=ADV say-PST.TR/ACT.1SUBJ soft=ADV=CONJ NEG
'I said ____ loudly not softly'

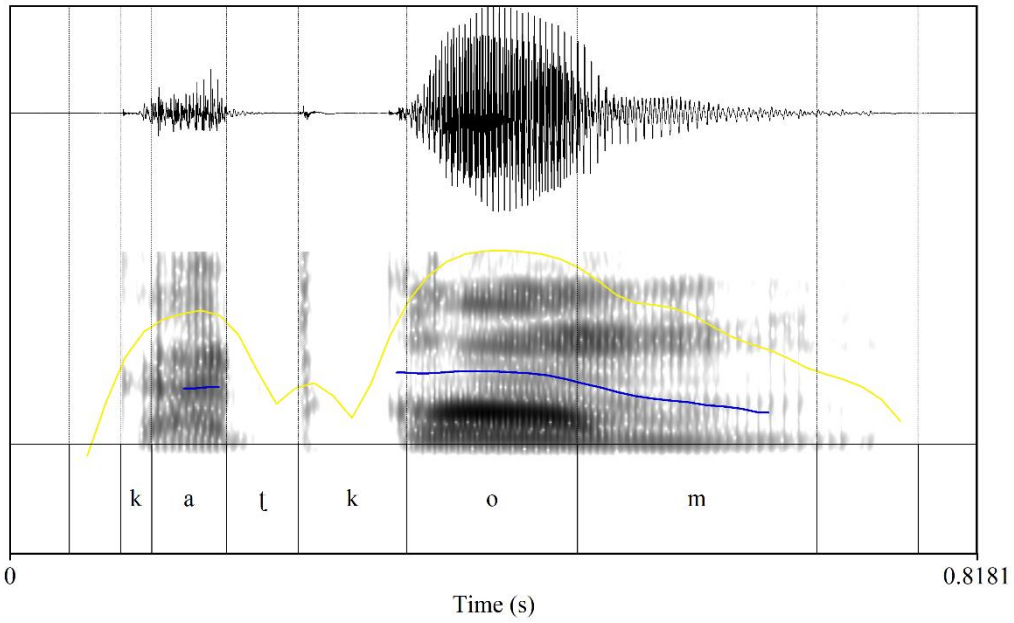
In both (1) and (2), the blank space is replaced by a target word from the text data. The phrasal position in (1) is intended to capture phrasal prominence in the target words and control for any speech perturbation that may be caused during the production of words in isolation, and the target position in (2) is intended to record the words in an out of focus or an unaccented intonational context to control for possible information structure effects. The fifty-one unique lexical items recorded thrice from six individuals produced a total sample size of 918 Santali disyllables, of which 46.13% have the CV.CV syllable structure with the vowels [a, i, e, o, u][2] as the syllable nucleus of both the first and second syllables. The data were recorded in a noise-free environment in the field using a head-worn unidirectional Shure mic connected by XLR cable to a Tascam linear PCM recorder, and the digital data are stored at a sampling frequency of 44.1 kHz and a bit depth of 32 bits in .WAV format.

2.2 Data analysis

The Santali speech data collected from the field were subjected to phonetic analysis by means of acoustic phonetic methods, for which purpose the data were manually annotated for word boundary and phoneme boundary in Praat (Boersma and Weenink 2020) using the spectral and temporal cues of speech sounds. Thereby, vowel sounds are annotated between the beginning and end of glottalic pulses, and sonorant sounds are annotated in low amplitude regions. In the case of the obstruents, they are annotated between the release of the oral closure and the beginning of the glottalic pulses in onset position and between the end of glottalic pulses and the point of oral closure in coda position (see Figure 2).

[2] The data set also includes three nasal vowels [ã, ĩ, ẽ] with [ã] appearing in the first syllable, [ĩ] appearing in the second syllable and [ẽ] appearing in both first and second syllables; a diphthong [ai] appearing in the second syllable and a lax front vowel [ɛ] appearing in the second syllable; they are for now not treated as unique syllable nuclei in the general analysis. However, for examining the interaction between syllable prominence and vowel types, the nasal vowels, the diphthong(s) and the lax front vowel are not included. Also, Ghosh (2008) includes schwa in his Santali vowel phoneme inventory, but Bodding (1922) considers it to be an allophone of [a] in words with high vowels. Until we have done instrumental analysis on the vowel system of various Santali lects, we reserve judgment on this issue too.

Figure 2: *Spectrographic illustration*[3] *of phonetic annotation of a Santali disyllable as produced by a male Santali speaker living in Sonitpur district of Assam*



2.2.1 Analyzing prominence in Santali disyllables

The analysis of prominence in Santali disyllables is based on three acoustic cues, namely, vowel duration, vowel intensity and fundamental frequency (Fry 1955; 1958). While vowel duration is calculated from the absolute length of the vowel nuclei in first and second syllables, vowel intensity is measured from the mean amplitude of the entire length of the vowel nuclei. Likewise, fundamental frequency is estimated from the mean of the entire vowel segment. Additionally, all values for the three acoustic parameters are normalized for speaker variability using the *z*-score normalization method *($z=x-\mu/\sigma$)*[4]. The normalized data is then analyzed and visually represented in R Version 3.5.3 (R Core Team, 2019) using the *ggplot2* package (Wickham et al. 2016) through its built-in functions of *geom_boxplot* and *geom_density* where the normalized values of vowel duration, vowel intensity and average fundamental frequency are treated as dynamic variables. Syllable positions and the contexts of the utterance (isolation, phrasal frame, unaccented frame) are treated as factor variables. Further, the same normalized data is tested with a one-way analysis of variance (ANOVA) using the *aov* function in R Version 3.5.3 (R Core Team 2019). For this purpose, normalized values of vowel duration, vowel intensity and average fundamental frequency are treated as dependent variables and syllable position (first versus second) and context of utterance (isolation, phrasal and unaccented) are treated as independent variables.

2.2.2 Analyzing segmental effect on prominence in Santali disyllables

This study also includes an investigation of segmental effects on syllable prominence in Santali disyllables. First, to examine the interaction between syllable prominence and vowel types, a subset of data, having only the five frequently occurring vowels [a, i, e, o, u] in first and second syllables of Santali disyllables, is used. In this analysis, vowel duration, vowel intensity and average fundamental frequency of the five vowels are measured separately in the two syllable positions. The same data is then visually represented using the built-in functions of *geom_boxplot* and *facet_grid* in the *ggplot2* package (Wickham et al. 2016) of R Version 3.5.3 (R Core Team, 2019), in which the normalized values

[3] This sound file can be heard and accessed in the Santali Living Dictionary at https://livingdictionaries.app/santali/entries/f0nG6dDN9GlqY0WPs6jC .

[4] z = Normalized value; x = Individually extracted values; μ = Mean of x; σ = Standard Deviation of x.

of vowel duration, vowel intensity and average fundamental frequency are treated as dynamic variables and syllable positions, vowel types and contexts of utterances are treated as factor variables.

Second, the effect of different consonantal environments on syllable prominence was examined. To do the analysis, a subset of data, including only the frequently occurring CV.CV syllable structure, that have onset consonants [h, k, m, r, tʃ] in both syllables, was used. Thereby, vowel duration, vowel intensity and average fundamental frequency in first and second syllables that have the five onset consonants were examined and compared with each other. This analysis is also represented using the same visualization method described above for examining the interaction between syllable prominence and vowel types except that the vowel type factor is changed to onset consonant type factor for this analysis.

3 Findings: Prominence in Santali disyllables

3.1 Vowel duration

Vowel duration is a common acoustic cue to determine prominence in languages whereby longer vowel duration is considered a robust indicator of prominence in languages of the world (Gordon and Roettger 2017). In the case of Santali (Ghosh 2008:23), vowel length is not reported to be phonemically distinct, and the data presented here exhibits only the phonetic length of vowel segments as produced in syllable nuclei of disyllabic words. By examining the speech data as produced by the Santali speakers recorded in this study, it was observed that vowel duration is generally longer in the second syllable as compared to vowel duration in the first syllable in Santali disyllables. Figure 3 demonstrates the vowel duration distinction in first and second syllables in the targeted disyllables as recorded in the three utterance contexts.

From Figure 3, it is evident that in Santali disyllables, average vowel duration in the second syllable is always longer than in the first syllable. Additionally, in Figure 4, the density plots reveal that the distribution of vowel duration in first and second syllables is distinct in all three utterance contexts. Also, a distinct skewing was observed in the data, indicating that longer vowel durations occur in second syllables, whereas shorter vowel durations occur in first syllables.

Figure 3: *Average Vowel Duration in First (1) and Second (2) syllables in Santali Disyllables.*

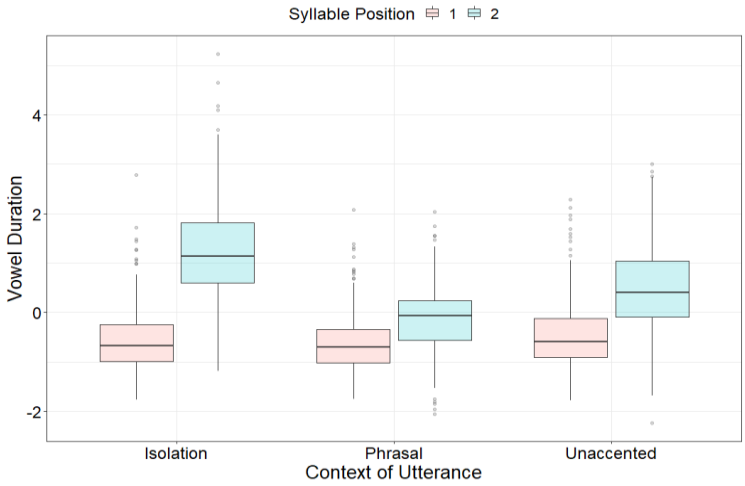
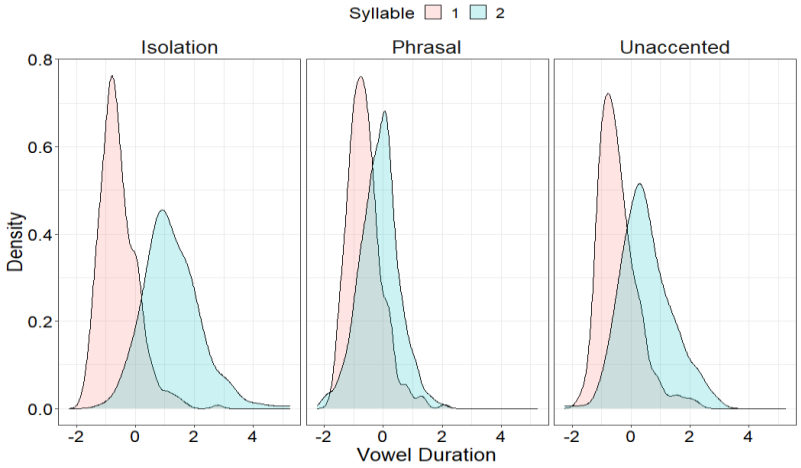


Figure 4: *Density distribution of vowel duration differences in first (1) and second (2) syllables in Santali disyllables.*



In this regard, it is notable that the distinction in vowel duration is greater when the disyllabic words are produced in isolation than when they are produced in the phrasal and in the unaccented intonational contexts.[5] Moreover, the difference in vowel duration in first and second syllables of Santali disyllables is found to be statistically significant in isolation [$F(1,642) = 796.4$, $p < 0.001$], in phrasal contexts [$F(1,629) = 101.3$, $p < 0.001$], and in unaccented intonational contexts [$F(1,622) = 248.9$, $p < 0.001$].

3.2 Vowel intensity

Vowel intensity refers to the acoustic energy in a vowel segment which is normally greater in prominent syllables than in non-prominent syllables. Unlike vowel duration, vowel intensity is a less robust cue for diagnosing prominence in languages, yet there is evidence that prominent syllables are distinct from non-prominent syllables with respect to their mean intensities (Gordon and Roettger 2017). Accordingly, in the present data, it is observed that both the averages as well as the overall distribution of average vowel intensity in first and second syllables of Santali disyllables are distinct from each other (see Figure 5 and 6).

From Figure 5, it is revealed that in Santali disyllables the average vowel intensity in second syllables is greater than the average vowel intensity in first syllables. Also, this is found to be true for Santali disyllables produced in all three utterance contexts. Likewise, the density plots in Figure 6 reveal that the distribution of average vowel intensity in first and second syllables is distinct wherein a skewing towards higher vowel intensity is observed in second syllables but a skewing towards lower vowel intensity is observed in first syllables. Also, the average vowel intensity difference in each of the three utterance contexts, namely, isolation [$F(1,642) = 30.62$, $p < 0.001$]; phrasal [$F(1,629) = 74.8$, $p < 0.001$] and unaccented intonational contexts [$F(1,622) = 33.72$, $p < 0.001$] are found to be statistically significant.

[5] This may indicate that vowel duration plays a role in demarcating the end of utterances, but this remains to be examined more systematically.

Figure 5: *Average vowel intensity in first (1) and second (2) syllables in Santali disyllables*

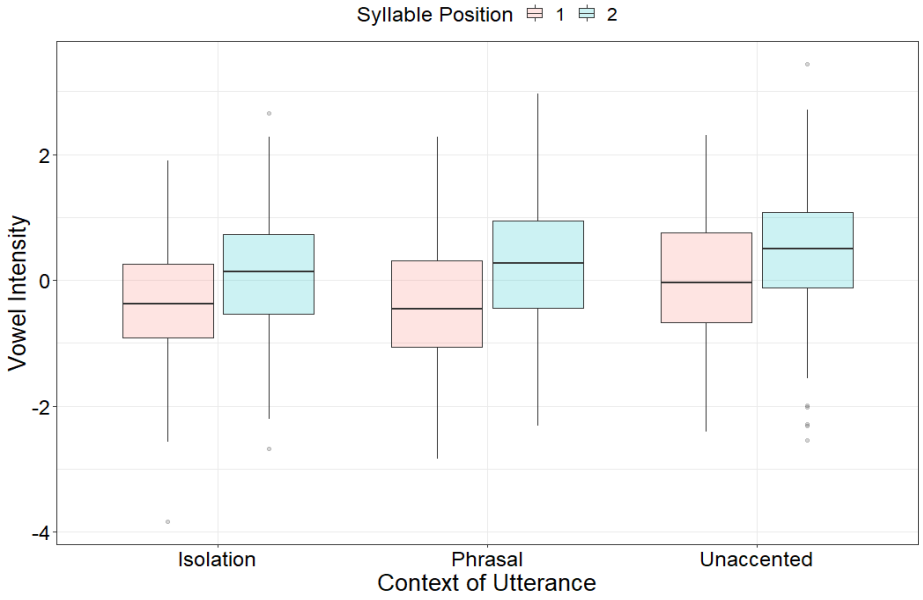
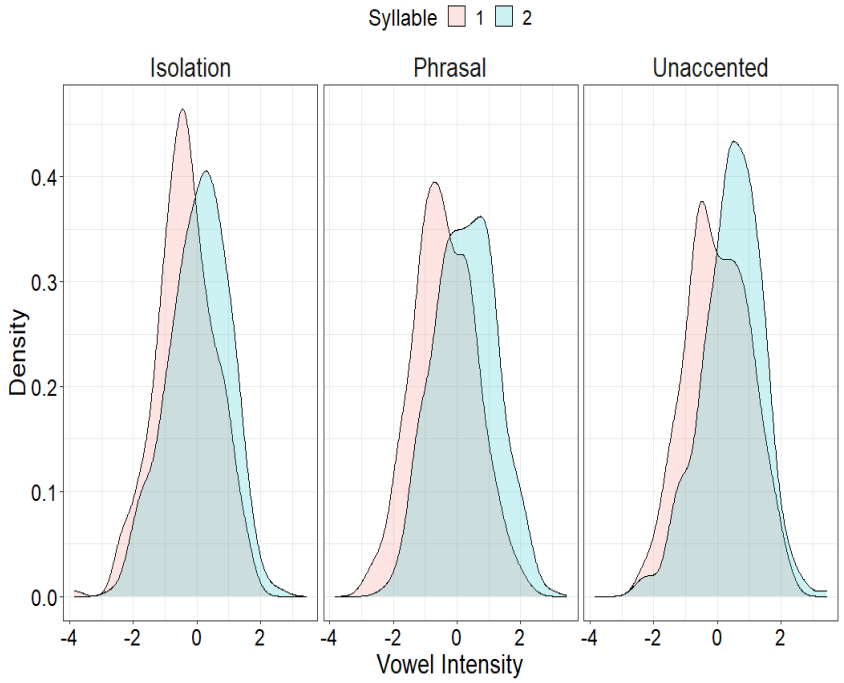


Figure 6*: Density distribution of vowel intensity differences in first (1) and second (2) syllables in Santali disyllables*



3.3 Fundamental frequency

Fundamental frequency represents pitch variation in speech sounds, and a systematic variation in pitch across syllables is known to be an indicator of syllable prominence in various languages (Gordon and Roettger 2017). Generally, prominence is associated with higher pitch which is expressed by greater fundamental frequency in the prominent syllable as opposed to lower fundamental frequency realized in the non-prominent syllable. The Santali data examined in this work reveals a similar pattern, but an exception is also observed in the analysis. Figure 7 shows the average fundamental frequency differences in first and second syllables of Santali disyllables.

Figure 7: *Average fundamental frequency in first (1) and second (2) syllables in Santali disyllables*

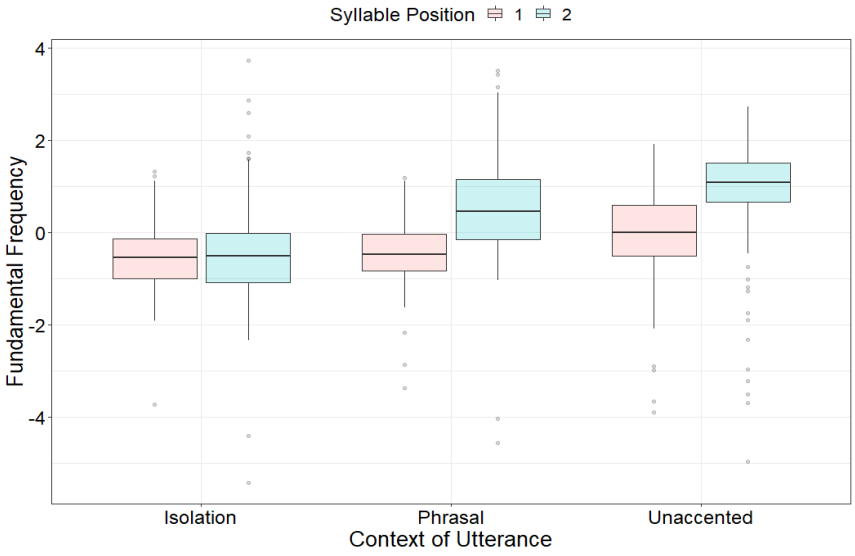


Figure 8: *Density distribution of Fundamental Frequency in First (1) and Second (2) syllables in Santali Disyllables*

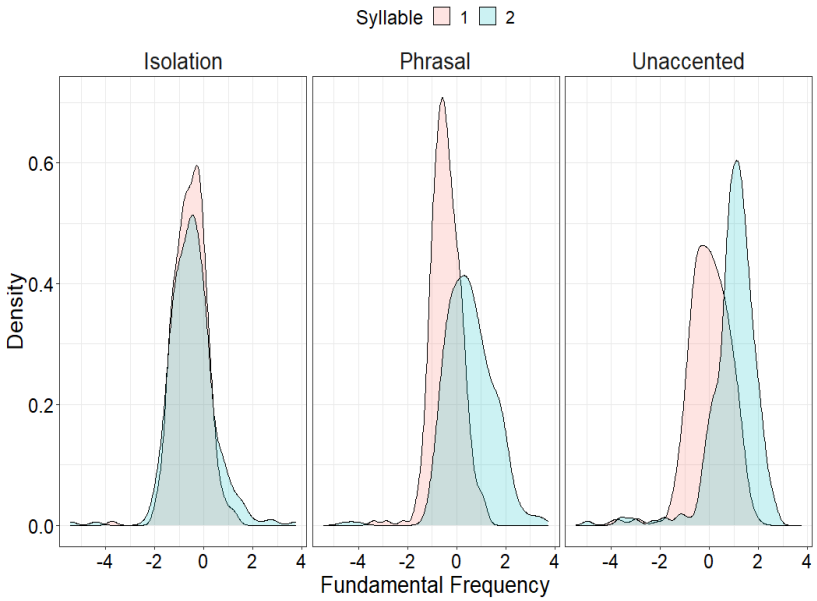


Figure 7 reveals that, in Santali disyllables, while the average fundamental frequency is higher in the second syllable than in the first syllable in the phrasal and unaccented intonational contexts, variation is absent when the disyllabic words are uttered in isolation. The same pattern is observed in the density plots presented in Figure 8, which shows that the distribution of average fundamental frequency in first and second syllables is not distinct in words that are spoken in isolation, whereas the same distributions appear to be distinct in words that are spoken in both the phrasal and unaccented intonational contexts. In this regard, the findings are also confirmed through statistical analysis whereby the average fundamental frequency difference in first and second syllables of Santali disyllables is found to be distinct with statistical significance when spoken in phrasal [$F(1,629) = 229.8$, $p < 0.001$] and unaccented [$F(1,622) = 183.1$, $p < 0.001$] intonational contexts but not when spoken in the isolation [$F(1,642) = 1.354$, $p = 0.245$] context. Thus, the analysis here suggests that although prominence, realized by higher pitch, in second syllable is present in Santali disyllables, the distinction is likely to be neutralized in words that are produced in isolation. However, this observation requires further

investigation[6] with substantial data evidence which could not be achieved in this preliminary study of prominence in Santali disyllables.

4 Findings: Segmental effects on syllable prominence

To explore the interaction between syllable prominence and segmental types and to see if the prominence of the second syllable over the first syllable in Santali disyllables is consistently maintained in different vowel types and in different consonantal environments, the three acoustic cues of prominence, namely, vowel duration, vowel intensity and fundamental frequency are further examined in this study using two separate subsets of data. Firstly, the five vowels in Santali [a, e, i, o, u], separated by their syllable positions, were examined to see if all vowel types render similar patterns of syllable prominence or not. Secondly, a subset of data containing only the CV.CV syllable structure and having the onset consonants [h, k, m, r, tʃ] in both syllables was examined to see if changes in the consonantal environment also effects syllable prominence or not.

4.1 Segmental effects on duration

With respect to vowel duration in first and second syllables having different vowel nuclei, it is observed that all vowel types have longer vowel duration only in the second syllable in all three contexts (see Figure 9). This implies that in Santali disyllables, the five vowels [a, e, i, o, u] are phonetically longer only when they occur in the second syllable but are relatively shorter when they occur in the first syllable, and that prominence of the second syllable manifested by longer vowel duration is not affected by changes in vowel types in the syllable nuclei. Likewise, by examining vowel duration in the environment of the five Santali consonants [h, k, m, r, tʃ], it is observed that in Santali disyllables that have any of the five consonants in the onset position, vowel duration is longer in the second syllable only, and not in the first syllable (see Figure 10). Also, the pattern is observed to be consistent in all three utterance contexts included in this study. This implies that in Santali disyllables, different onset consonants do not have an impact on syllable prominence cued by vowel duration.

[6] A further investigation of pitch variation in first and second syllables can be achieved by extracting pitch values (fundamental frequencies) at different pitch timings including initial, medial, and final. This can show potential pitch changes that may be undisclosed in the average pitch distinction in the two syllable positions in the isolation context. Also, maximum, and minimum pitch values can be extracted, which can be utilized for examining potential separation of pitch ranges in the two syllable positions when the disyllables are produced in isolation.

Figure 9: *Average vowel duration in first (1) and second (2) syllables in Santali disyllables having different vowel nucleus*

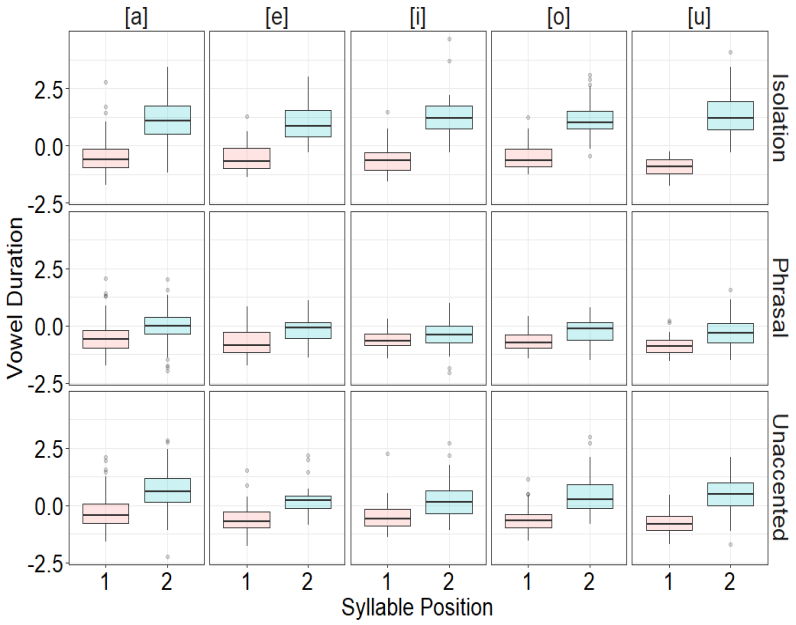
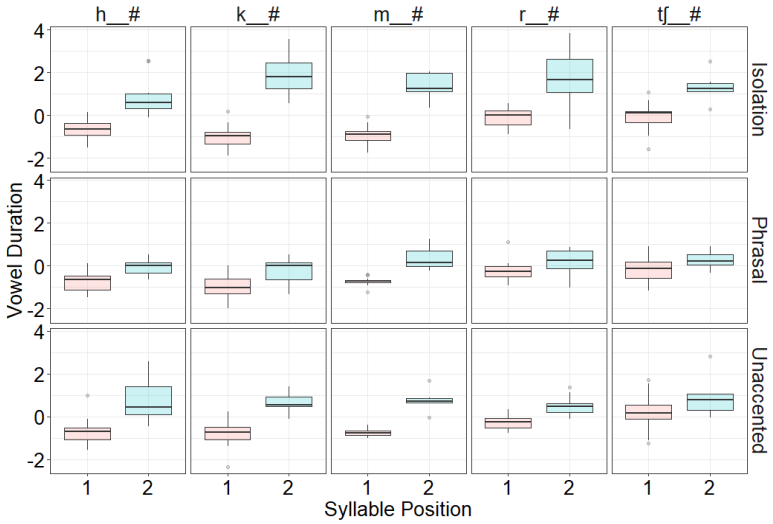


Figure 10: *Average vowel duration in First (1) and Second (2) syllables in Santali disyllables with CV.CV syllable structure having different consonantal environments*



Thus, based on the data presented in Figures 9 and 10, it is confirmed that vowel duration is a robust cue for identifying prominence in Santali disyllables whereby neither the vowel types nor the onset consonants in CV.CV syllable structures affect the manifestation of prominence in the second syllable.

4.2 Segmental effects on intensity

The general examination of vowel intensity in first and second syllables of Santali disyllables revealed that vowel intensity is normally higher in the second syllable. In the micro analysis, while the tendency to produce higher intensity in the second syllable is found consistently in the mid vowels [e, o] and the low vowel [a], in the case of the two high vowels [i] and [u], the vowel intensity difference between the two syllables is observed to be neutralized and even reversed for the high front vowel [i] in the unaccented intonation context (see Figure 11). This indicates that the two high vowels [i] and [u] in

Santali disyllables do not have distinct vowel intensities between the two syllables. Also, while the distinction is observed to be equally neutralized in all three utterance contexts for the back high vowel [u], in case of the front high vowel [i], the difference is neutralized in isolation and in the phrasal context but is reversed in the unaccented intonational context. Thus, from analysing the interaction between vowel types and vowel intensity in Santali disyllables, it is revealed that prominence of second syllable cued by higher vowel intensity is maintained in disyllables that have either the mid vowel or a low vowel in the syllable nuclei, whereas the pattern is likely to be neutralized or reversed if the disyllables have only the high vowels in their syllable nuclei.

Figure 11: *Average vowel intensity in First (1) and Second (2) syllables in Santali disyllables having different vowel nucleus*

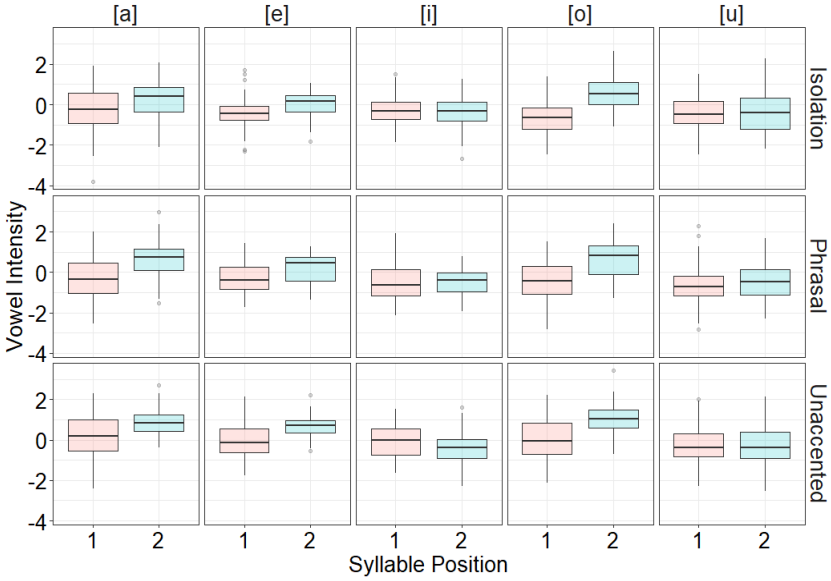
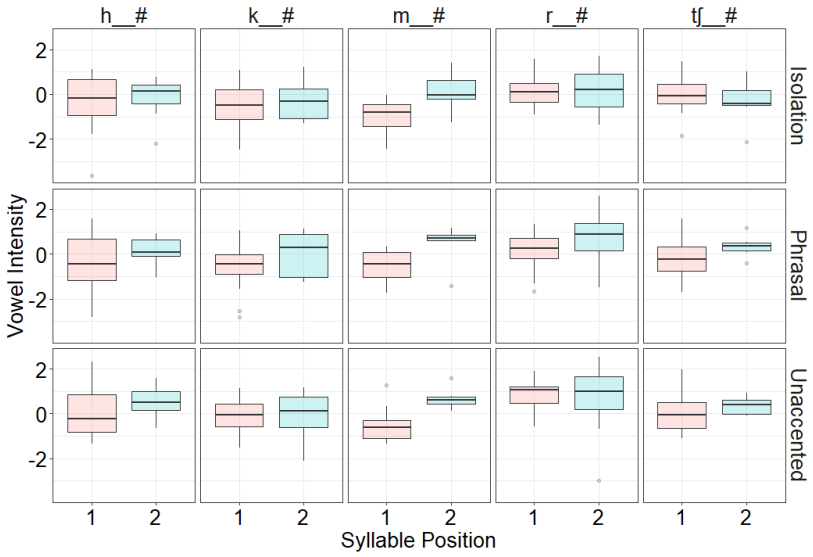


Figure 12: *Average vowel intensity in First (1) and Second (2) syllables in Santali disyllables with CV.CV syllable structure having different consonantal environments*



In addition to the variations in vowel intensity arising from vowel types, an examination of the interaction between vowel intensity and different onset consonants reveals that vowel intensity differences in first and second syllables may be neutralized or reversed even in the environment of

certain consonant types in the onset position. Of the five onset consonants examined in this study, the analysis here reveals that the general tendency to produce higher vowel intensity in the second syllable has only been consistently maintained in all three contexts of utterances in the environment of the bilabial nasal consonant [m] (See Figure 12). On the other hand, vowel intensity distinctions between first and second syllables are observed to be neutralized in the environment of the voiceless velar consonant [k] and alveolar trill consonant [r] in the isolation and unaccented intonational contexts. The difference is seen to be reversed in the environment of the voiceless palatal affricate consonant [tʃ] in the isolation context only. These observations imply that like vowel types, onset consonant types in the CV.CV syllable structure also have an impact on the rendering of vowel intensity in Santali disyllables, whereby consonants such as [m] and [h] appear to have no or minimal impact. In contrast, consonants such as [k] and [r] may neutralize the intensity distinction, and consonants such as [tʃ] may even reverse the vowel intensity distinction in Santali disyllables.

Thus, from analysing the interaction between vowel intensity and vowel types and onset consonant types it is revealed that, unlike vowel duration, vowel intensity is a less robust cue for identifying syllable prominence in Santali disyllables. Specifically, at the micro level, it is found that there are certain segmental exceptions that probably suppress the general tendency to produce higher intensity vowels in the second syllable of Santali disyllables.

4.3 Segmental effects on fundamental frequency

Fundamental frequency difference in Santali disyllables has been shown to be sensitive to various utterance contexts whereby higher fundamental frequency in the second syllable is exhibited only in the phrasal and unaccented intonational contexts but not in the isolation context of utterance. Significantly, the same pattern of fundamental frequency distinction is observed when the five Santali vowels [a, e, i, o, u] are analysed separately with respect to their syllable positions in the disyllables. Figure 13 presents the vowel-wise fundamental frequency distinction in first and second syllables of Santali disyllables in the three contexts of utterances included in this study. From Figure 13 it is evident that all five Santali vowels have higher fundamental frequency in the second syllable when they are produced in the phrasal and unaccented intonational contexts, but as an exception not in the isolation context. This implies that differences in vowel types do not impact the second syllable prominence depicted by higher fundamental frequency in Santali disyllables.

Similarly, by examining the effect of different onset consonants on the fundamental frequency of the vowel nuclei of Santali disyllables, it is observed that besides the lack of fundamental frequency distinction in first and second syllables in the isolation context of utterance there is only a minimal impact of onset consonant types even in the phrasal and unaccented intonational contexts of utterances. Figure 14 presents the fundamental frequency difference in first and second syllables of Santali disyllables that are grouped according to their onset consonants [h, k, m, r, tʃ] and the contexts of utterances that are included in this study.

Figure 13: *Average fundamental frequency in first (1) and second (2) syllables in Santali disyllables having different vowel nucleus*

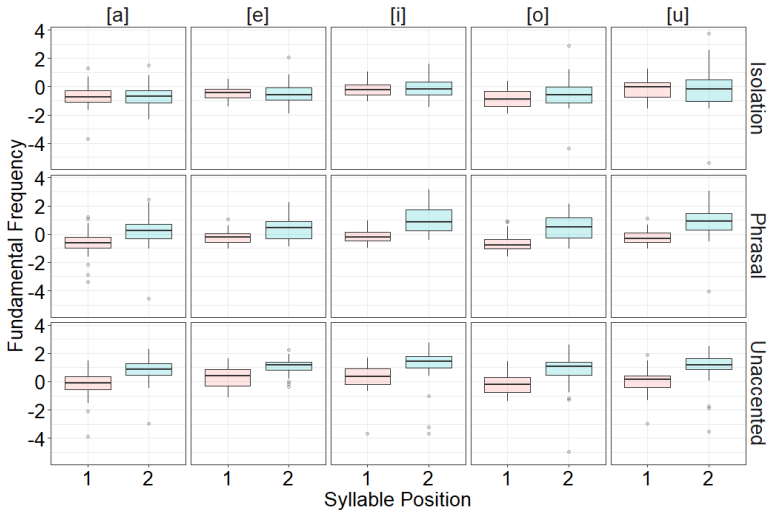
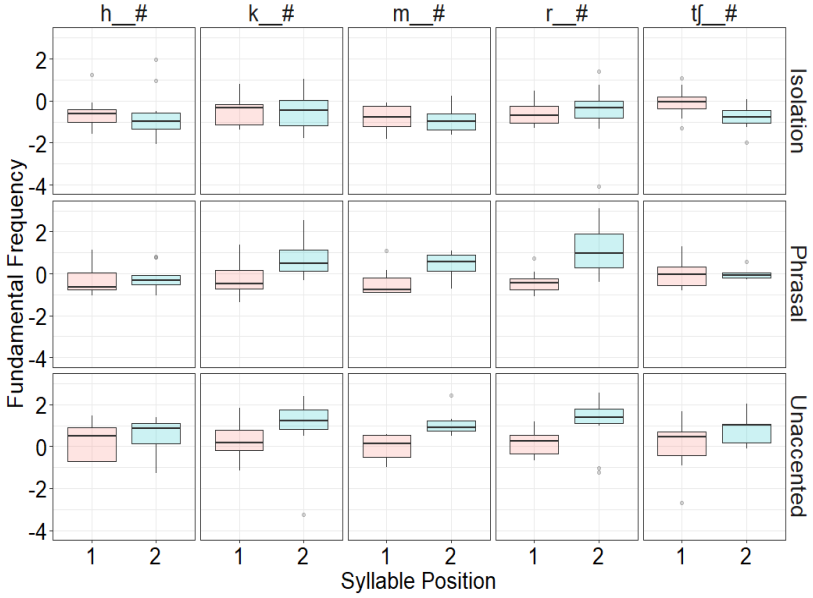


Figure 14: *Average fundamental frequency in first (1) and second (2) syllables in Santali disyllables with CV.CV syllable structure having different consonantal environments*



From Figure 14, it is observed that while a fundamental frequency difference between first and second syllables of Santali disyllables may remain neutralized in different consonantal environments when the target word is said in isolation for words beginning in [k, m, r], in the very same context, the difference may even be reversed in the environment of the glottal fricative consonant [h] and voiceless affricate consonant [tʃ]. Also, the same consonantal environments appear to have resulted in neutralizing or minimizing the fundamental frequency difference between first and second syllables of Santali disyllables even when they are produced in the phrasal and unaccented intonational contexts of utterances.

Thus, the analysis of vowel-wise fundamental frequency differences in first and second syllables of Santali disyllables revealed that vowel types do not impact syllable prominence in the language. Additionally, the exceptional case of neutralizing fundamental frequency differences in first and second syllables in the isolation context of utterance is also confirmed from the examination of the interaction of vowel types and fundamental frequency in Santali disyllables. However, an analysis of

different onset consonants indicates that two consonant types, namely, fricatives and affricates may either minimize the fundamental frequency difference in first and second syllables or even reverse the difference when Santali disyllables with CV.CV syllable structures bearing the two consonants in the onset position are recorded in isolation only.

5 Comparison with Sora

Despite claims in the literature to the contrary (Donegan 1993, Donegan and Stampe 1983, 2004), instrumental analyses show that Sora clearly has second syllable prominence in disyllabic forms (Horo and Sarmah 2015, Horo 2017, Horo, Sarmah and Anderson 2020). Phonetic data representing Sora speech varieties of four geographical locations in Assam, namely, Singrijhan, Sessa, Lamabari and Koilamari and one geographical location in Odisha, namely, Raiguda, provide evidence that vowels in the second syllable are longer, louder and pitched higher than the vowels in the first syllable of disyllables. Accordingly, the acoustic cues of prominence reveal that vowel duration, vowel intensity and fundamental frequency (*f0*) are generally higher in second syllables only; see Figures 15 to 17.

Figure 15: *Average vowel duration in first and second syllable of Sora disyllables (Horo 2017)*

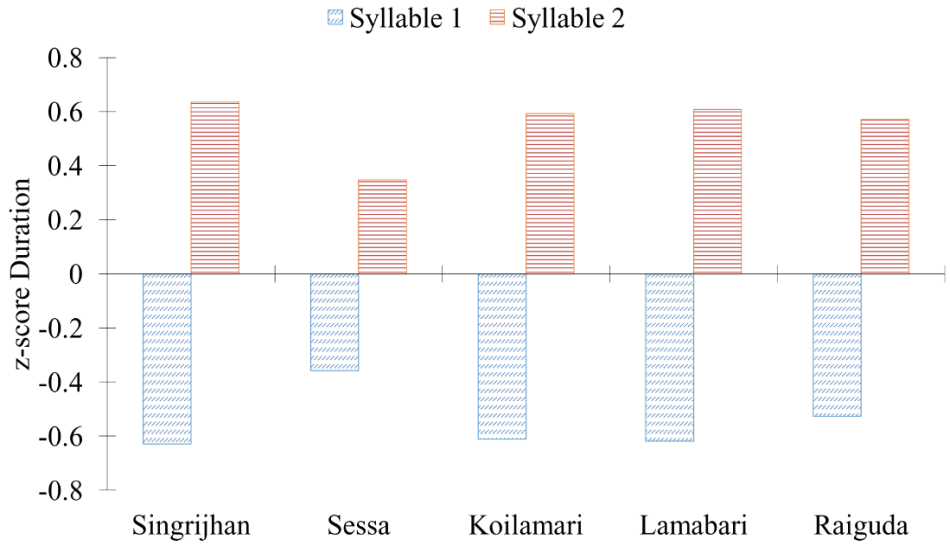


Figure 16: *Average f0 in first and second syllable of Sora disyllables. (Horo 2017)*

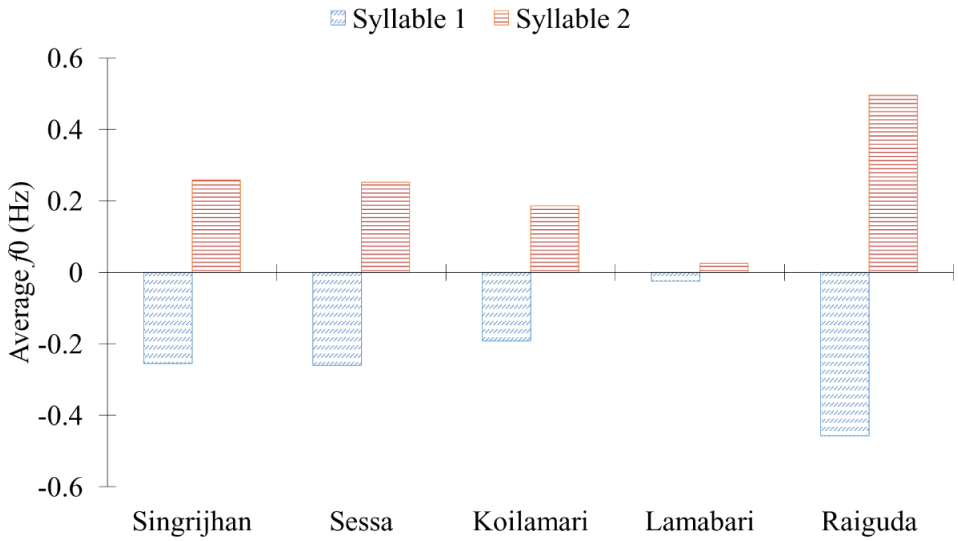
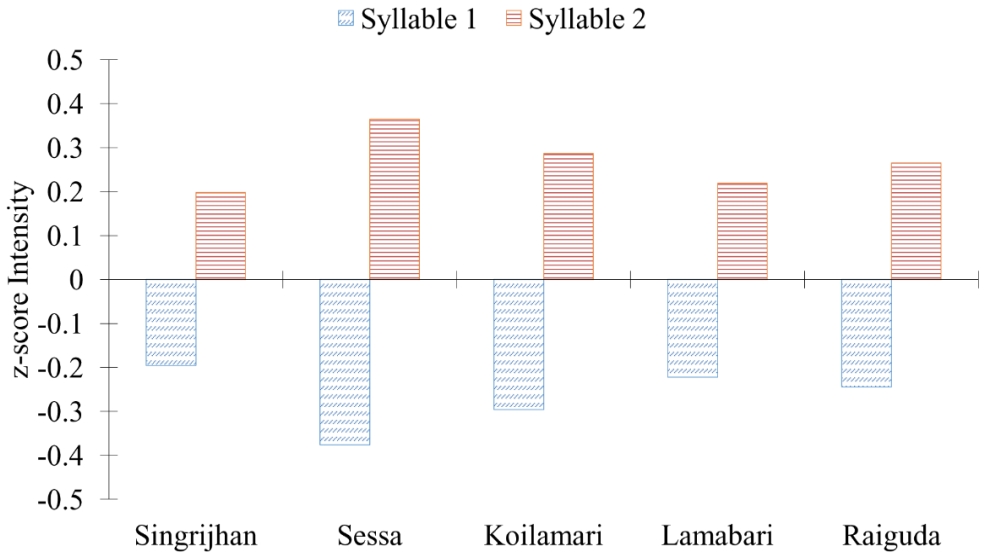


Figure 17: *Average vowel intensity in first and second syllable of Sora disyllables. (Horo 2017)*



Likewise, other Munda languages have been claimed to show similar patterning as well (e.g., Remo (Bhattacharya 1968) or West Bengal Santali (Ghosh 2008)), but no instrumental data has been offered in support of this, even while we believe those published observations to be accurate. However, Assam Santali, based on preliminary phonetics analysis and supported by acoustic evidence, shows that second syllable prominence is the pattern attested in disyllabic words.

On the other hand, Indo-Aryan and Dravidian languages, including those of Odisha, at least those studied instrumentally to date which may in fact be limited to Oriya (Mahanta 2010) and Telugu (which is not an official language of Odisha but is spoken in Parlakhemundi in the extreme south of Odisha, on the border) typically are trochaic or first-syllable prominent in similar contexts having more peripheral vowels in the first syllable and more phonological contrasts attested, all pointing to first-syllable prominence (Khan 2016).

6 Discussion

The phonetic analysis supported by acoustic evidence above suggests that duration and intensity are largely consistent cues of prominence in the Assam lect of Santali when investigated across all three utterance contexts, viz., in isolation, in a phrasal frame and in an unaccented frame. In each instance, the second syllable reflects greater prominence with respect to these cues than does the corresponding first syllable in these disyllabic words. Fundamental frequency is also distinct and points to prominence on the second syllable over the first syllable in Assam Santali as well. However, unlike duration and intensity, fundamental frequency is only statistically relevant as an acoustic cue of prominence in phrasal and unaccented utterance contexts. In the isolation context on the other hand, fundamental frequency does not cue word-level prominence *per se* in the Assam Santali lect examined. Simultaneously, as the isolation context may potentially reflect word-level intonation as well as intonation of the full utterance level–two areas where pitch variation can well be distinct–there may be intonational parameters that operate on the full utterance level that interact with and potentially override word-level intonational or prominence cueing parameters. However, what these may be requires further research. Also, duration does not appear to show sensitivity to specific vowel types or consonant contexts in their functions as cues of prominence. Intensity differences between first and second syllables on the other hand appear to be sensitive to both the presence of high vowels and the presence of [k], [r] or [tʃ] in onset position, while fundamental frequency may show some sensitivity to consonantal environment, specifically the presence of an initial fricative [h] and affricate [tʃ].

Furthermore, that the phrasal frame data in our study–which is a potentially inherently a quasi-focus position–and the unaccented position frame–which is explicitly an out of focus–largely pattern

together with respect to the three acoustic cues of prominence suggest that focal intonation is not active *per se* in determining the distribution of the three examined cues of word-level prominence.

7 Broader Munda perspectives

That our data show that the Assam Santali lect under investigation showing second syllable prominence in disyllables may be surprising to some scholars, but others will note this might be expected due to previous analyses of West Bengal and Jharkhand Santali lects. Those who may find this surprising might be scholars who would either predict that Santali should be similar to data from other South Asian language groups such as the Indo-Aryan lects that it is in contact with, or they may assume this based on claims that have been made in previous investigations, not of Santali specifically, but hypotheses that are said to apply to all Munda languages, and thus Santali by extension.

To be sure, one camp of scholarship has long asserted that Munda languages exclusively and only show trochaic prosodic patterns (i.e., first syllable prominence in disyllables). This is best and most succinctly encapsulated in the theory of 'rhythmic holism' put forth by Donegan (1993) and Donegan and Stampe (1983, 2004). They suggest that there was a one-time shift from iambic 'rhythm' to trochaic rhythm in Munda at the proto-Munda level, and that it was this prosodic shift that triggered a series of cascading changes that caused the wholescale typological restructuring of proto-Munda from isolating to agglutinative-synthetic and to verb-final syntactic structure, and so on, which resulted in modern Munda languages ultimately representing the mirror image of their Austroasiatic sister languages that remained in Southeast Asia. However, such claims have been refuted by various scholars in recent work on Munda, such as Horo and Sarmah (2015), Horo (2017a), Horo et al. (2020), Anderson (2015b), Anderson (2020) and Ring and Anderson (2018).

As mentioned above, other scholars have made claims about the prosodic structure of Santali in print that do not align with our experimental findings, although it should be mentioned here that no acoustic or statistical data have been offered by any of the other scholars mentioned below with respect to the Santali data, so their claims must therefore be considered impressionistic and preliminary as a result. Thus, Neukom (2001:8) claims that in disyllabic stems, in Santali, "(s)tress falls on the first syllable; however, if the first syllable is light and the second heavy (iambic structure), stress falls on the second syllable". Therefore, he suggests that Santali reflects a Quantity Sensitive system. Our data as presented above on disyllables do not support this view. Regardless of syllable shapes in the words, the pattern is always the same: the acoustic cues that indicate prominence in Santali disyllables, including intensity, fundamental frequency, and duration, conspire to make second syllables prominent in Santali disyllables over the first syllable in such words. This is also true regardless of whether the word appears in a quasi-focal frame or in an explicitly out of focus frame for all three examined cues of prominence, and in isolation as well for duration and intensity.

Note as also mentioned above that not all previous researchers agree with Neukom's take on the Santali data. Thus, Ghosh (2008:30) states "(s)tress is always on the second syllable of the word regardless of whether it is an open or a closed syllable". Our instrumental acoustic and statistical data do in fact support Ghosh's view, at least in uninflected disyllables.

As the present investigation is just a preliminary study, we have for the time being limited ourselves to only examining disyllabic lexemes in this one lect, Assam Santali. Bodding's (1922) musings on the topic of the system of prominence attested in Santali are rather involved. In short, casting things into a modern typology of prosodic systems or prominence assignment, he suggests that when taking into consideration all Santali words, including inflected forms of verbs which can create rather lengthy morphological strings (or g[rammatical] words), the system of prominence should be considered to be morpholexically specified and constrained. He suggests that there may be certain morphemes capable of bearing stress while others do not. Whether this will be verified instrumentally remains a task for future research.

It may also be the case that there are mismatches between phono-prosodically defined word domains (such as prominence or vowel harmony, see Anderson, Horo and Harrison 2022) and larger morphological complexes (g-words) that may constitute more than one phono-prosodic 'word' or

represent phrases prosodically, even if functioning syntactically as units.[7] Moreover, it appears that in Munda languages more generally there are morphological word elements or grammatical morphemes that appear to be more integrated into the phono-prosodic domain of roots and form single phono-prosodic words with such lexical hosts, but others remain outside of such word domains. Put differently, some grammaticalized elements may be affixes and others rather clitics,[8] some eligible to be assigned prominence for example and some that are not, and yet others that are variable in this regard. A task for our future research on Santali is to determine what the prosodic and morphotactic characteristics of the full range of inflectional elements are that may constitute Santali morphological word complexes and how exactly these interact (or do not interact) with the system of prominence assignment. Only once such a study is complete will we be able to definitively determine the entire system of prominence that characterizes this important Kherwarian Munda language.

Santali is far from alone within Munda in presenting a confusing picture asserted in different publications about what the system of prominence might be. A wide variety of other, often conflicting, claims about the intonational structure of individual Munda languages have appeared in print. With few exceptions (e.g., Rehberg 2003 for Kharia) instrumental data are not used as the basis for the analysis offered, so the interpretations remain largely impressionistic. With respect to Rehberg (2003), she proposed that low pitch on the initial syllable in a disyllabic word of Kharia followed by a high pitch is what signalled prominence, which while of course possible, seems largely motivated by a desire to conform the attested data in Kharia with the standard view that Munda languages are trochaic, such that it must be low pitch that signals prominence in Kharia if this trochaic pattern is true. We reserve further consideration of the Kharia data until we have had a chance to subject the data to our own instrumental analysis, but simply comment that the language seems to have a five-vowel system of phonemes and the realization of the name of the language has a schwa in the initial and allegedly prominent syllable, even while the schwa does not appear to be a phoneme in the language. Furthermore, intensity may increase concomitant with the raising of pitch in second "unaccented" syllables in Kharia (Peterson 2011) further underscoring that Kharia likely has second syllable prominence in disyllables, not first-syllable prominence.

Mundari, a language closely related to Santali, embodies the lack of clarity about the system of prominence that typifies our present understanding of Munda languages as expressed in print. Even whether 'stress' exists per se is debated, as Osada (1992:36) considers Mundari to be a pitch accent language, while Cook (1965:100), Langendoen (1963:14-15), and Sinha (1975:39), consider Mundari to be a stress language. But these latter three researchers do not agree on what the system of stress is. Sinha considers the language to have a quantity sensitive system whereby disyllabic words of the shape $C^1V^1C^2V^2$ or $C^1V^1C^2V^2C^3$ (where C^2 can also be a homo-organic nasal+stop sequence) stress the *second* syllable, but in disyllabic words of the shape $C^1V^1C^2C^3V^2$, stress falls on the initial syllable and in trisyllabic words, stress falls on the 2nd syllable regardless of the shape. Cook (1965) states that only if the final syllable is closed, it is accented, otherwise it is the initial syllable in disyllabic words. Osada (2008:104) states that if a word is trisyllabic, stress can *only* be on the second or the third syllable: on the third syllable if that is not a suffix, otherwise it falls on the second syllable in Mundari trisyllabic words, but *never* on the first syllable, regardless of syllable weight.

[7] This was suggested about Munda and Khasian language to explain some observed differences between apparent phrases vs. lexemes in Khasian Pnar and Sora (Ring and Anderson 2018).

[8] Santali clearly reflects such a system with subject markers which function as clitics in both imperfective and perfective series of inflections. Subject clitics in Santali do not even preferentially target the verbal 'word' as their host, but rather target the word immediately preceding the verb, or they may appear at the end of the morphological verbal word (or on occasion both places simultaneously). See Anderson (2007, 2015a, 2015b, 2020) for more details.

Table 2: *Munda language prominence patterns and acoustic cues (based on Hildebrandt and Anderson 2021)*

Language	Prominent Syllable	Acoustic Cues/Notes
Sora	Second syllable in disyllables (Horo 2017)	Duration, Intensity, Fundamental Frequency
Gorum	Penult/final syllable (Anderson & Rau 2008)	Not yet formally investigated
Gutob	Final/second syllable (Voß p.c.)	H-Pitch, other cues unclear. Not yet formally investigated
Remo	Second syllable (Bhattacharya 1968)	Intensity, Pitch, Duration? Not yet formally investigated
Gtaʔ	Second/final syllable (Anderson in preparation)	Not yet formally investigated
Kharia	Conflicting: Initial/second/final? (Rehberg 2003, Peterson 2011)	L-Pitch (Initial) H-Pitch (Second/final) Intensity on non-initial syllables. Not yet formally investigated
Juang	Conflicting: Second vs. initial (Patnaik 2008, Dasgupta 1978)	Not yet formally investigated
Koɽowa	Second/last syllable of stem (Barker 1953)	Duration? Not yet formally investigated
Ho	Conflicting: QS/initial or Morpholexical (Nottrott 1882, Pucilowski 2013)	Not yet formally investigated
Mundari	Conflicting: Pitch accent vs. Stress accent Conflicting: Initial vs second vs. Morpholexical (Osada 1992, 2008; Cook 1965; Sinha 1975; Hoffmann 2001; Langendoen 1963)	Pitch, other cues? Not yet formally investigated
Santali	Conflicting: Second syllable vs. QS/initial (Bodding 1922, Neukom 2001, Ghosh 2008)	Intensity, Duration, Pitch to some degree
Korku	Second/QS? (Zide 2008)	Not yet formally investigated

However, all this aside, our data clearly shows that in disyllables, prominence is found on the second syllable in the Assam Santali lect we discuss here. Therefore, Santali appears to reflect the same pattern that has previously been identified for both Assam and Odisha lects of the distantly related Sora language, also of the Munda family (Horo and Sarmah 2015, Horo 2017a, Horo et al. 2020). Given the typological, geographic and genetic distance between Santali and Sora, and given strong areal tendencies against second syllable prominence in South Asia, and the fact that most of language groups that Munda is related to phylogenetically within the Austroasiatic phylum show a similar (and seemingly) cognate system of prominence assignment, one might be tempted to suggest that these data may point us in the direction of assuming that proto-Munda may well have been second syllable prominent in disyllables as the simplest explanation for the observed parallels between Sora and Santali, and impressionistically, other Munda languages as well, such as Gtaʔ, in addition to their demonstrable similarity to other Austroasiatic groups.

8 Future Research Goals: Towards an intonational typology of the Munda languages

In this paper we are beginning the first step on a long journey to compare the interface between phonological structure and prosodic features and the morphosyntax of the Munda languages. First, we must determine what the patterns and cues of prominence for each of the languages are, or at least a representative set of the languages. From basic uninflected words, we expand this typology for inflected

forms of words and phrases and see what patterns emerge and how these different elements combine and whether such combinations exhibit distinct phono-prosodic patterns. This larger work is underway for Sora currently and will expand to Santali in the next year. We will extend this similar investigation to other Munda languages representing different branches of the family tree with an eye to not only grounding future discussions of Munda prosodic structures in instrumental phonetics analysis supported by acoustic evidence, but also how these structures interface with the complex morphosyntax of the languages.

9 Summary

None of the previously mentioned scholars relied on instrumental data for their analyses, which therefore remain impressionistic. Our study is based on instrumental analysis and suggests that Santali (at least as spoken in Assam) *always* shows prominence on the second syllable of disyllables, cued by intensity, f_0 and duration. This suggests that Assam Santali shows a pattern similar to that attested in Assam and Odisha lects of Sora (Horo and Sarmah 2015, Horo 2017, Horo et al. 2020), contra Donegan and Stampe (2004). In other words, it is an iambic pattern and not quantity sensitive, at least in disyllables.

However, resolving what possible historical development the Munda languages have undergone both individually and at the proto-language level remains several steps away as we must first engage in the systematic synchronic analysis of word, phrase, and utterance level prominence systems (and subsets thereof) for the attested languages of the family before we have an adequate empirical basis to engage in such far reaching but important questions. The present study is just a preliminary step on this journey.

References

Anderson, Gregory D. S. 2007. *The Munda Verb. Typological Perspectives*. Berlin: Mouton de Gruyter.

Anderson, Gregory D. S. 2015a. Overview of the Munda languages. In M. Jenny and P. Sidwell (eds.) *The Handbook of Austroasiatic Languages*. Volume 1. Grammar Sketches of the World's Languages. Mainland and Insular South East Asia. Amsterdam: Brill. 364-414.

Anderson, Gregory D. S. 2015b. Prosody, phonological domains and the structure of roots, stems and words in the Munda languages in a comparative/historical light. *Journal of South Asian Languages and Linguistics* 2 (2): 163-183.

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

Anderson, Gregory D. S. in preparation. The Gtaʔ Language. *Plains Gtaʔ and Hill Gtaʔ Texts, Lexicon and Grammar.*

Anderson, Gregory D. S., Luke Horo and K. David Harrison. 2022. Vowel Harmony in the Munda Languages. Submitted to H. van der Hulst and N. Ritter (eds.) *Oxford Handbook of Vowel Harmony*. Oxford: OUP.

Anderson, Gregory D. S. and Felix Rau. 2008. Gorum. In Gregory D. S. Anderson (ed.) *The Munda Languages.* Routledge Language Family Series. London: Routledge (Taylor and Francis), 381-433.

Barker, Phillip R. 1953?/nd. *The phonemes of Korowa*. Unpublished manuscript. Seattle, Washington.

Bhattacharya, Sudhibhushan. 1968. *A Bonda Dictionary*. Building Centenary and Silver Jubilee Series 18. Poona [Pune]: Deccan College.

Bodding, P. O. 1922. *Materials for a Santali Grammar, mostly phonetic*. Dumka: Santal Mission.

Cook, Walter A. 1965. *A Descriptive Analysis of Mundari: A study of the structure of the Mundari language according to the methods of linguistic science.* Ph.D. Dissertation, Georgetown University.

Dasgupta, Dipankar. 1978. *Linguistic Studies studies in Juang, Kharia Thar, Lodha Mal-Pahariya, Ghatoali Pahariya.* Calcutta: Anthropological Survey.

Donegan, Patricia J. 1993. Rhythm and vocalic drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman Area* 16.1:1-43.

Donegan, Patricia J., & Stampe, David. 1983. Rhythm and the Holistic Organization of Language Structure. In J. F. Richardson, M. L. Marks, & A. Chukerman (eds.), *Papers from the parasession on the Interplay of Phonology, Morphology and Syntax*. Chicago Linguistic Society, Chicago: 337-353.

Donegan, Patricia J., & Stampe, David. 2004. Rhythm and the synthetic drift of Munda. In R. Singh (ed.), *The Yearbook of South Asian Languages and Linguistics*. Mouton de Gruyter, Berlin, Germany: 3-36.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America 27*, no. 4: 765-768.

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and speech 1*, no. 2: 126-152.

Ghosh, Arun. 2008. Santali. In Gregory D. S. Anderson (ed.) *The Munda Languages*. Routledge Language Family Series. London: Routledge (Taylor and Francis), 11-98.

Gordon, Matthew K. and Timo. B. Roettger 2017. *Acoustic Correlates of Word Stress: A Cross-Linguistic Survey*, Linguistics Vanguard, 3(1). DOI: https://doi.org/10.1515/lingvan-2017-0007. Accessed 4 October, 2021.

Hildebrandt, Kristine A. and Gregory D. S. Anderson. 2021. Word Prominence in Languages of Southern Asia. To appear in Ksenia Bogomolets and Harry van der Hulst (eds.), *Word prominence in polysynthetic languages*. Oxford: Oxford University Press.

Horo, Luke. 2017. *A Phonetic Description of Assam Sora*. Ph.D. Dissertation, Indian Institute of Technology, Guwahati.

Horo, Luke & Priyankoo Sarmah. 2015. Acoustic analysis of vowels in Assam Sora. In L. Konnerth et al. (eds.), *Northeast Indian Linguistics* 7:69-86. Canberra: ANU.

Horo, Luke, Priyankoo Sarmah and G. D. S. Anderson. 2020. Acoustic Phonetic Study of the Sora Vowel System. *Journal of the Acoustical Society of America* 147.4:3000-3011.

Langendoen, Terry. 1963. *Mundari Phonology*. Unpublished manuscript. Cambridge, Massachusetts.

Neukom, Lukas. 2001. Santali. *Languages of the World/Materials* 323. München: Lincom.

Nottrott, Rev. A. 1882. *Grammatik der Kolh-Sprache.* Gütersloh: C. Bertelsmann.

Osada, Toshiki. 1992. *A reference grammar of Mundari*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Osada, Toshiki. 2008. Mundari. In Gregory D. S. Anderson (ed.), *The Munda Languages*. Routledge Language Family Series. London: Routledge (Taylor and Francis), 99-164.

Patnaik, Manideepa. 2008. Juang. In Gregory D. S. Anderson (ed.) ,*The Munda Languages*. Routledge Language Family Series. London: Routledge (Taylor and Francis), 508-556.

Peterson, John M. 2011. *Grammar of Kharia*. Leiden: Brill.

Pucilowski, Anna. 2013. *Ho Morphology and Morphosyntax*. Ph.D. Dissertation, University of Oregon.

Registrar General of India. 2011. *Distribution of the 100 Non-Scheduled Languages – India, States & Union Territories – 2011 Census*. New Delhi: Office of the Registrar General and Census Commissioner, India.

Rehberg, Kerstin. 2003. *Phonologie des Kharia–Prosodische Strukturen und segmentales Inventar*. MA Thesis, Universität Osnabrück.

Sinha, N. K. 1975. *Mundari Grammar*. Mysore: Central Institute of Indian Languages.

Wickham H., Chang W., Wickham M.H. 2016. *Package 'ggplot2'*. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2: 1-189.

Zide, Norman H. 2008. Korku. In Gregory D.S. Anderson (ed.), *The Munda Languages*. Routledge Language Family Series. London/New York: Routledge (Taylor and Francis), 256-298.