Towards a typology of information structuring in Munda languages

Gregory D. S. Anderson, *Living Tongues Institute for Endangered Languages*
Bikram Jora, *Living Tongues Institute for Endangered Languages*
Moses Oyeleye, *Living Tongues Institute for Endangered Languages*
Izza Armand, *Living Tongues Institute for Endangered Languages*
Tara Shabazaz, *Living Tongues Institute for Endangered Languages*
Luke Horo, *Living Tongues Institute for Endangered Languages*

Abstract

Munda languages use morpholexical, syntactic and prosodic means to encode information structure of various types. These parameters are independent but may be interdependent. Based on two distantly related languages of the family, Sora of the Sora-Gorum branch and Mundari of the Kherwarian branch we provide details on each of these three means of encoding information structure. There is some evidence that two of the morpholexical means that are productive in Mundari and restricted in Sora may be old formants encoding information structure in the family. Prosodic means appear to show no correlation across the languages and syntactic means of encoding information structure show some commonalities that could be retentions of an earlier proto-language system, while others are clearly language-specific.

1 Introduction

In this paper we discuss formal means of encoding information structure in two Munda languages, Sora of the Sora-Gorum subgroup and Mundari of the Kherwarian subgroup. We discuss three logically independent but often interconnected means of highlighting discursively and pragmatically important arguments, predicates and adjuncts in these languages, viz., morpholexical indices §2, syntactic configurations §3 and prosodic means §4 to track or highlight elements that are topical(ized) or in focus.

Under morpholexical means of encoding focus or topic, we include clitics, affixes or prosodically full forms of functional words. Under syntactic configurational means of encoding information structure, we include scrambled syntax or the pragmatically and information structurally activated orders of elements. Cleft-structures seem to be an alien strategy to Munda, as no such structures occur in our collection of spontaneous discourse texts. Under prosodic means we cover the differential use of the acoustic cues of prominence of duration, intensity or pitch/f0 to encode not lexical or grammatical or word-level or phrasal or utterance level contrasts but rather discourse/pragmatic and information structural ones. This is the first ever analysis of the different formal means of encoding information structure in the Munda language family.

In descriptions of Munda languages, the concepts of topic and focus get confused or conflated and the morphological elements involved may indeed straddle both functions.

This functional overlap is particularly true in contrastive contexts where contrastive topic and contrastive focus may converge. To be sure, even in non-contrastive contexts, both topic and focus reference arguments and events in a discourse space larger than a sentence, shifting attention and noteworthiness or referentiality of such referents and events (and indeed adverbial adjuncts) as suits the narrative or conversational structure. Of course, topic is an accepted node or functional projection in Chomskyan syntax and has a formal role in Lexical Functional Grammar in the functional domain. Nevertheless, some scholars dispute topic as a singular concept but rather view it as a cluster of prototypes (Asher, 2004; Jacobs, 2001).

In traditional approaches, topic should represent given information and focus conversely should be new information on some meta-discursive level. But topicalization can reintroduce a referent for example that may have dropped from the immediate attention of the hearer. In many formal syntactic approaches, and indeed supported by cross-linguistic evidence, topics may be external to the rest of the sentence and occupy a separate field in the syntax often on the left-periphery. But focalized elements can also appear in such a position, which further contributes to the potential confusion between these two information structuring concepts. Munda languages support this generalization as exemplified in relevant places below.

Topic is widely used in linguistic analysis even while defining it is elusive (Asher, 2004), and indeed while the concepts of contrastive topic, discourse topic and sentential topic are interconnected, they formally have little in common other than a vague sense of 'aboutness', and being 'known' information.

With respect to focus some standardly followed or accepted definitions in linguistics include those of Lambrecht (1994) and Dik (1997). Lambrecht defines focus as "the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition. The focus component is by definition an unpredictable part of the proposition." Dik (1997) for his part uses the following definition "[t]he focal information in a linguistic expression is that information which is relatively the most important or salient in the given communicative setting and considered by S[peaker] to be the most essential for A[ddressee] to integrate into his pragmatic information."

Many of the other nuanced takes on this complex and somewhat elusive phenomenon of 'focus' include those found in such a diverse array of studies as works by Bearth (1999), Büring (1997, 2010), Caron (2000), Chafe (1976), Dik (1980, 1981), Dik et al. (1981), Drubig & Schaffar (2001), Fiedler et al. (2010), Gundel & Fretheim (2004), Krifka (2004, 2007), Sasse (1987), Uhmann (1991), van der Wal (2016), Zimmermann (2008, 2016), Zimmermann & Onea (2011) and Zubizarreta (1998), all of which we consulted in an attempt to determine the functional domains of the morpholexical elements used in Sora and Mundari that structure information in spontaneous narrative and conversational discourse.

Like topic, there are several different meanings of focus as used in the literature, whether the communicative point in is informational, presentational, contrastive, corrective or counter-presuppositional. The scope of the focalization phenomena also varies, including whether the focus is on an argument or is predicate/event-centered, or whether the focalizing is ex-

haustive, restrictive, additive, emphatic, verum-based, etc., semantically.

2 Morpholexical means of encoding information structure in Munda

Use of morpholexical means in flagging focus and topic is the most common strategy found in Munda languages. This includes uninflected particles, affixes and/or clitics used to encode information structuring semantics.

2.1 Morpholexical (-cum-syntactic) Means of Encoding Focus in Sora

Contrastive topic in Sora is marked by the postposing of the particle *ati*/*ate* or the clitic/affix *-ti*/*te* after the topicalized element. It sometimes occurs left-dislocated in a quasi-cleft-like field on the left periphery of the clause.

(1) a. ao kən a-kəɽibab-ən ate titti-n sed-ir-ai-dʒi
DISC this DEP-skull-N.SFX TOP over.there-N.SFX throw-go-PFV-3PL
'so as for that skull of his, they took and threw it away over there'

b. bəru-n a-dʒaʔmol ati daku biɲdo
mountain-N.SFX DEP-seed TOP COP but
saroba-n a-dʒaʔmol ati ɲen arika
field-N.SFX DEP-seed TOP I NEG.COP
'I have some mountain seeds, but as for field seeds, I don't have any'

The most common focalizing element in Sora is *=na*. It is quite promiscuous in the host it selects for (2a-2f). It commonly is translated as 'really, surely, indeed'.

(2) a. Verb
a-pəsiʔ ati anin kəjiʔ-te=na kəjiʔ-ti
DEP-child TOP 3PRON die-NPST=FOC die-NPST
'as for the child, he will die, he will definitely die'
(note also the left dislocated TOP–marked NP)

b. Object NP
kun a-kansim-ən=na kudub-ən tubob-tə-dʒi
that DEP-rooster-N.SFX=FOC all-N.SFX kill(head)-NPST-PL
'they kill all the roosters (that were brought)'

c. Manner Adverb
ɲen enegoi=na doʔoŋ-bin sum-le ambin dʒanaŋ ɲen
1SG like.this=FOC body-2PL catch-CV 2PL also 1SG
kabɲid-ta-bin
kill-NPST-2PL.UND
'like this I will catch you, and I will kill you also'

d. Postposition

kon baʈ̪i=na anin-dʒi a-rəb-rəb-ɖ̪om-li-n-dʒi anin-dʒi
that with=FOC 3PRON-PL DEP- protect-RFLXV-ITR-N.SFX-PL 3PRON-PL

sənu-ba-n dʒanaŋ jir-rə-dʒi
fight-DIR-N.SFX also go-PST-PL

'they would go also to fight (battles/wars) for their protection with these'

e. Temporal adverb/adjunct

adnəŋ=na kəm pasidʒ-an kəna-n paŋ-ʈ̪e
at.that.time=FOC this child-N.SFX tiger-N.SFX take-NPST

a-monsil-ɲen ɡam-le
DEP-niece-1SG say-PST

'they would go also to fight (battles/wars) for their protection with these'

f. Subject/Agent

o-ɲen=na ajer-təm ɡam-li
child-1SG.POSS=FOC return-PRF say-PST

'my child really has returned, she said'

With locative adverbs/adjuncts, the focalizer =*na* has an exhaustive/restrictive focal interpretation in Sora.

(3) a. tidtin=na
there:N.SFX=FOC
'only there'

b. iʔdʒa ʈ̪iʔnen=na
no here:N.SFX=FOC
'no, only here (her uncle said)'

2.2 Morpholexical Means of Encoding Focus in Mundari

There are at least two affixes/clitics - that function in the information structure system in Kherwarian languages: -*do* and –ɡ*e*. In Mundari -*do* appears to be a topicalizer (4) mainly while -ɡ*e*/-ɡ*i* functions in a manner similar to Sora =*na*, with similar promiscuity in host selection (5a-5d).

(4) Mundari *do*

enlo=do miad hoɽo hente-te lolo lidʒa
that.moment=TOP one man there-ABL hot cloth

tusiŋ-kete oɖ̪oŋ-dʒa-n-a
wear-CV come.out-TAM-ITR/MDL-IND

'at that very moment one man came out from there wearing a hot cloth'

(5) a. Subject

aɲ-ge=ɲ senog-a
1SG-FOC=1SG.SUBJ go:IPFV-IND

'I will go' (Osada, 1992, p. 136)

b. Object

manɖi-ge=ɲ dʒom-ke-n-a
food-FOC=1SG.SUBJ eat-AOR-ITR/MDL-IND

'I ate the food' (Osada, 1992, p. 136)

c. Locational/Directional NP

Ranci-ate-ge=m hidʒuʔ-aka-n-a
Ranchi-ABL-FOC=2SG.SUBJ come-ANT-ITR/MDL-IND

'you have come from Ranchi' (Osada, 1992, p. 137)

d. Verb

manɖi=ko dʒom-tan-ge-a
food=3PL.SUBJ eat-IPFV-FOC-IND

'they are eating food' (Osada, 1992, p. 137)

3 Syntactic Means of Encoding Focus in Munda languages

Text data suggests SOV is overwhelmingly dominant in all Munda languages. SVO order on the other hand focuses on the subject.

(6) Sora

mənra-n-dʒi gijʔ-le-dʒi ənsəlo-n-dʒi=adoʔoŋ
person-N.SFX-PL see-PST-PL woman-N.SFX-PL=OBJ

'*the men* saw the women'

(7) Mundari

sumit dal-dʒa-d-ko-a=e hon-ko
Sumit beat-TAM-TR/ACT-3PL.OBJ-IND=3SG.SUBJ child-PL

'*Sumit* beat the children'

Flipping objects and subjects focuses on the object and does so exhaustively in Sora or adds an emphatic connotation 'indeed'.

(8) a. Sora

ənsəlo-n-dʒi=adoʔoŋ mənra-n-dʒi gijʔ-le-dʒi
woman-N.SFX-PL=OBJ person-N.SFX-PL see-PST-PL

'the men saw *only the women*'

b. Sora

bogad sonum-ən amaŋ anlen pur-re-bi-n=den
if spirit-N.SFX DEP:OBJ.II we worship-PST-1PL-N.SFX=COND

'if indeed it is the spirit we have worshipped'

In Mundari, this scrambled word order rather also focalizes subjects.

(9) Mundari

hon-ko sumit dal-dʒa-d-ko-a=e
child-PL Sumit beat-TAM-TR/ACT-3PL.OBJ-IND=3SG.SUBJ

'*Sumit* beat the children'

OVS order in Sora marks exhaustive focus on the object but also carries a passive connotation.

(10) a. Sora

ənsəlo-n-dʒi=adoʔoŋ gijʔ-le-dʒi mənra-n-dʒi
woman-N.SFX-PL=OBJ see-PST-PL person-N.SFX-PL

'*only the women* were seen by Opino'

b. Sora

anin-dʒi sanil-ta-dʒi gordʒaŋ-mər-ən-dʒi
3PRON-PL follow/escort-NPST-PL village-person-N.SFX-PL

'they were followed by the villagers'

OVS order in Mundari on the other hand conveys focus on the object. Basically any element that appears post-verbally in Mundari typically assigns a focalized reading on that element.

(11) Mundari

hon-ko=e dal-dʒa-d-ko-a sumit
child-PL=3SG.SUBJ beat-TAM-TR/ACT-3PL.OBJ-IND Sumit

'Sumit beat the *children*'

VSO order in Sora focalizes the subject. This formation is somewhat common in our large text corpus.

(12) Sora

aŋanʈi dʒʰundʒum-boʲdʒ-ən a-ɖinna jər-ai-ʈen
when see.off-bride-N.SFX DEP-day come-CLOC-PRF

karra-n sinʈa-n-eʈin kən-i məndra
excess-N.SFX worry-ITR/MDL-PRF this-FOC man

‘When the day came for the sending off of the bride this very man was extremely worried.’

The word that immediately follows an initial verb in Mundari is in contrastive focus (13), which of course can include a subject NP (14), with the result being VSO order.

(13) Mundari

dal-dʒa-d-ko-a=e sumit hon-ko
beat-TAM-TR/ACT-3PL.OBJ-IND=3SG.SUBJ Sumit child-PL

‘*Sumit* beat the children.’

VOS in Sora is reported to focus the whole VP exhaustively according to speaker interviews, but it is textually unattested in our collection of sixty texts in six varieties of Sora.

(14) Sora

dʒum-le kinte-n opino-n
eat-PST banana-N.SFX Opino-N.SFX

‘Opino *already surely ate the banana only*’

VOS order entails focus on the object in Mundari on the other hand, as the element immediately following an initial verb focalizes it, including the object NP.

(15) Mundari

dal-dʒa-d-ko-a=e hon-ko sumit
beat-TAM-TR/ACT-3PL.OBJ-IND=3SG.SUBJ child-PL Sumit

‘Sumit beat *the children*.’

Cleft constructions, whether of the Wh- or ‘it’-type do not seem to form part of the natural repertoire of devices used for focalization in either Sora or Mundari, or at least no such formations spontaneously occur in our text corpus and are rejected by speakers when generated.

4 Prosodic Means of Encoding Illocutionary Force (and Focus?) in Munda languages

The acoustic cues associated with phonetic prominence in the Munda languages include vowel intensity, vowel duration, and fundamental frequency, e.g., in Sora (Horo et al., 2020, 2022, 2023a; Horo & Anderson, 2021a) or Assam Santali (Horo & Anderson, 2021b, 2023a). However, there is no commonality across all the languages as to which cue or sets of cues is favored for word- or phrase-level prominence and which may be used to indicate other categories, e.g., word boundary, phrasal edge or information structure.

For example, in Sora one finds vowel intensity peaks on the second syllable followed by a lengthening of word-final syllables to mark word-edge. Phrasally, pitch declines at the end except in Y/N questions where it rises.

4.1 Sora prosody, illocutionary force and focus

In Sora declaratives, with no special information structural properties, there is a decline in pitch in each word and the lowest pitch is on the final syllable of the utterance; see Figure 1 'little by little I will get cold' [∼little cold feel-NPST-1SG.UND].

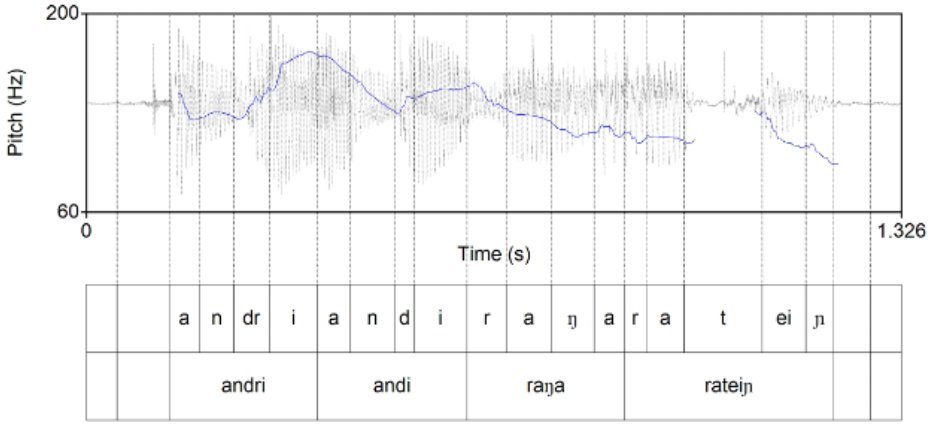

Figure 1: Pitch contour in a declarative sentence in Sora

In imperative formations as in Figure 2 'sleep carefully!' [one self-CV-I/M-sleep-I/M-IMP], all but the final word show pitch rising throughout the word and the imperative verb ending taking low pitch.

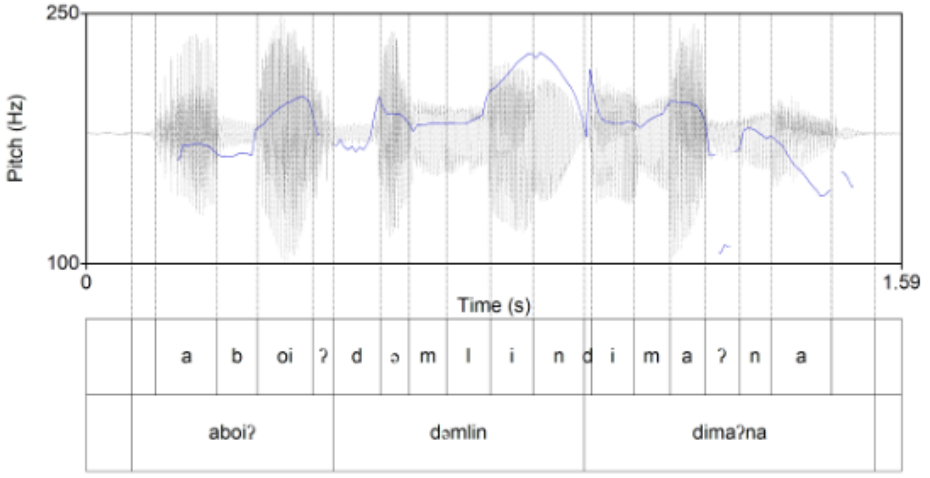

Figure 2: Pitch contour in an imperative sentence in Sora

Interrogative sentences with Wh-words show a similar pattern; see Figure 3 for 'what did you say?' [2SG what say-PST]. The Wh-element is by definition under focus and correspondingly there is high pitch on this element.

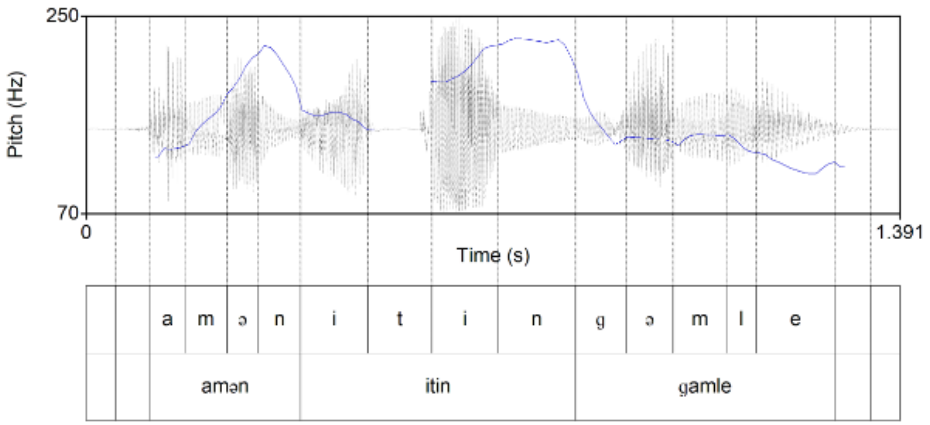

Figure 3: Pitch contour in a wh-word interrogative sentence in Sora

Yes/No questions, however, have rising pitch on the utterance final word in Sora which here is the interrogative particle *po*, as in Figure 4 for 'are you voiceless?' [2SG NEG.NMLZR-VOICE Q]. The entire predicate is in focus in such constructions and consequently there is also a rise in pitch on the Q-particle as is quite common cross linguistically, but noteworthy in Sora as there is typically a declination in pitch at the end of utterances including Wh-questions (see Figure 3) and declaratives (Figure 2).

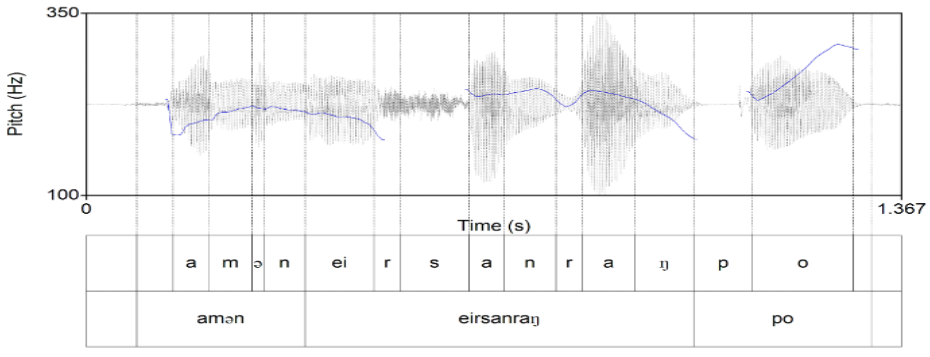

Figure 4: Pitch contour in a yes/no interrogative sentence in Sora

However, forced interpretation of elements as focalized in Sora (not using the focalizing element *=na* nor syntactic scrambling), using Odia and English translation to reinforce the

focalization shows no consistent realization acoustically in Sora, e.g., whether contrastive focus is on the subject or the predicate, these show the same pitch contours; see Figures 5 and 6.

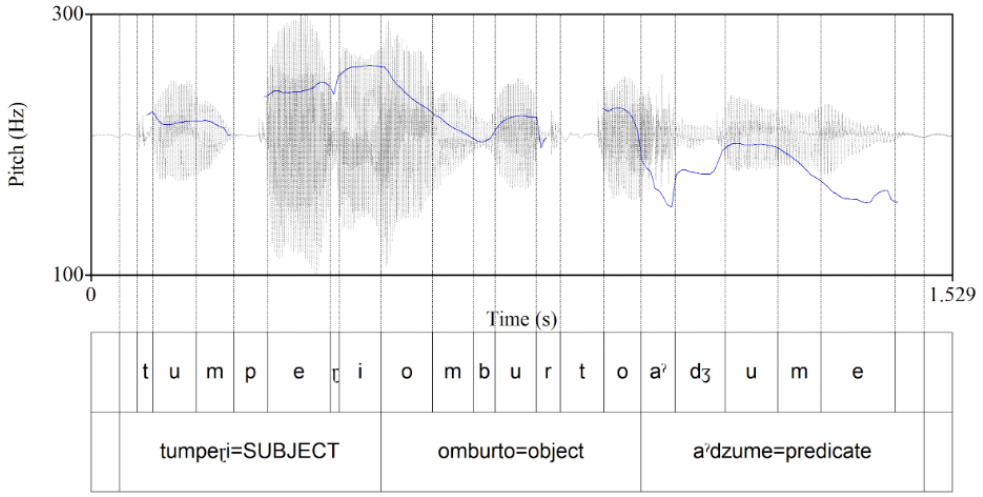

Figure 5: Pitch contour in a declarative form with contrastive focus on subject

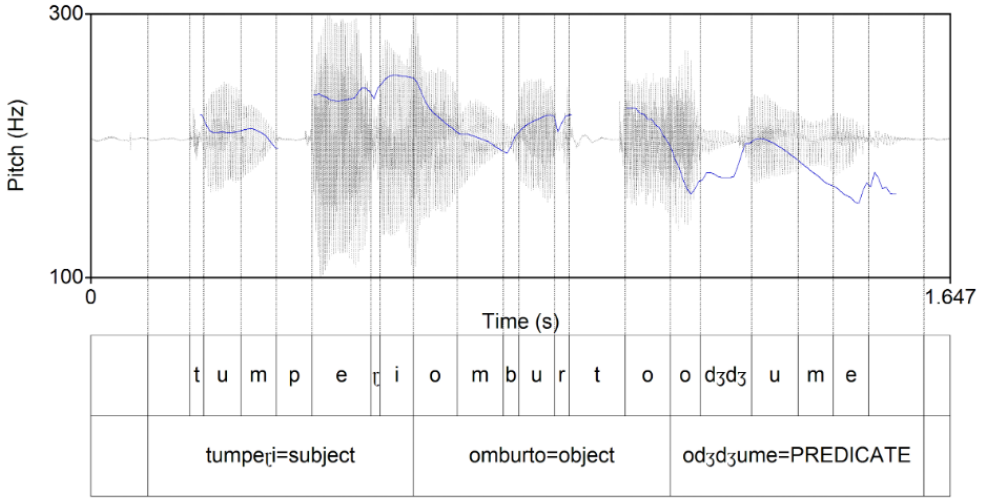

Figure 6: Pitch contour in a declarative form with contrastive focus on predicate

Therefore, prosodic means alone do not seem to be able to be used to encode focus in Sora, rather this should also accompany morpholexical means such as the interrogative particle *=po* or *=na*.

4.2 Mundari prosody and focus

Mundari disyllables show no consistent acoustic cue of lexical prominence. Longer duration on the second syllable is found in isolation (Figure 7), while intensity (Figure 8) and pitch (Figure 9) rise across the word. Note that all Mundari recordings were made in four contexts: in isolation (i), in a sentential/phrasal frame (p), a non-focal frame (n), and in a focal (f) frame.

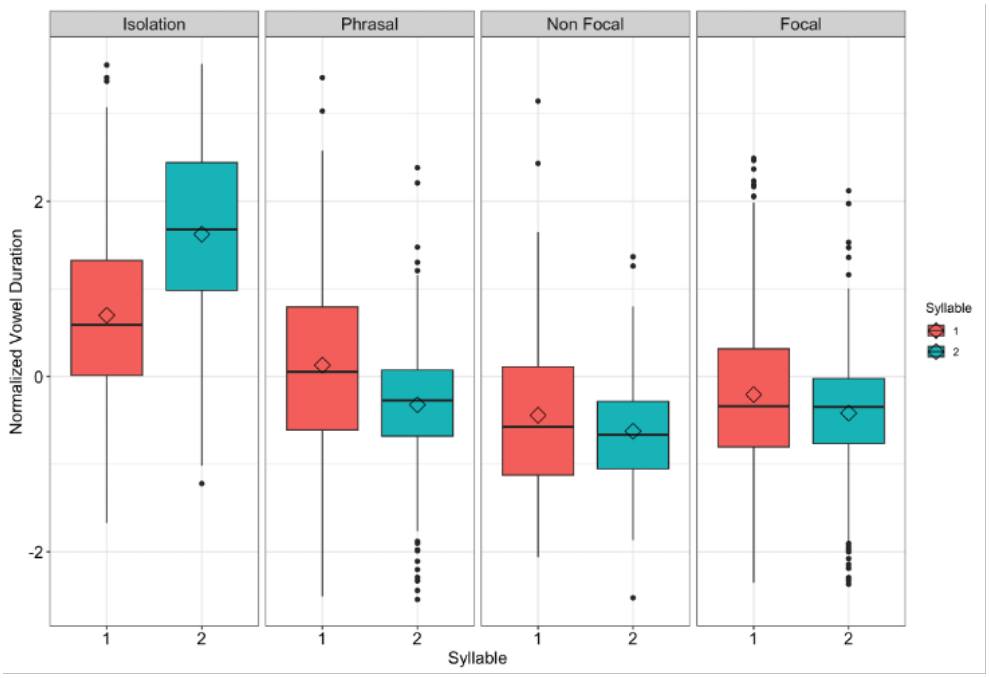

Figure 7: Distribution of vowel duration in Mundari disyllables produced in isolation and non-isolation contexts

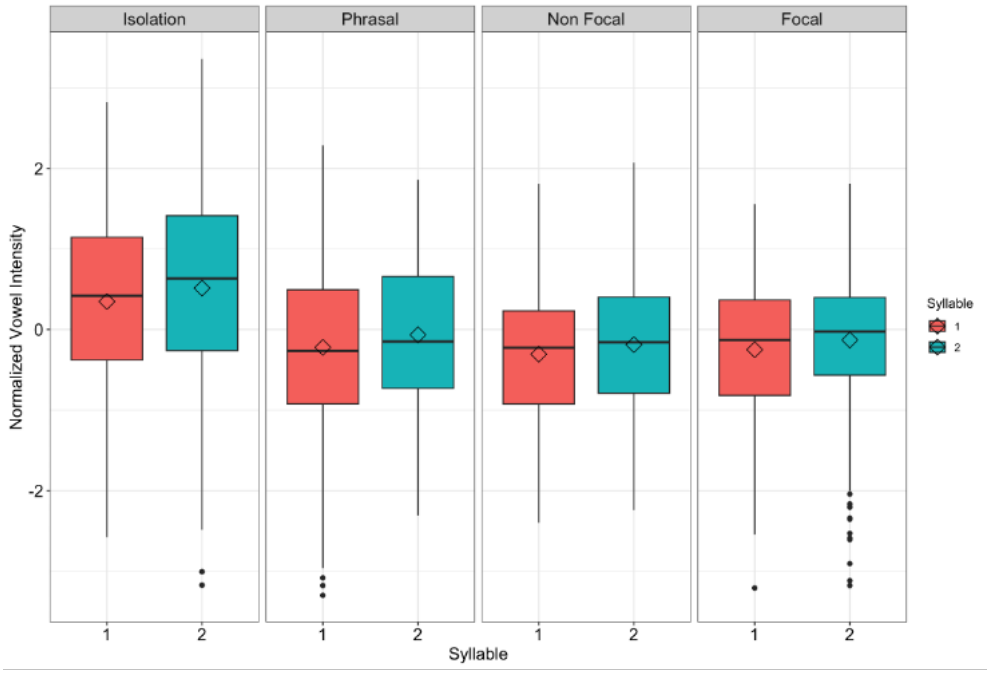

Figure 8: Distribution of vowel intensity in Mundari disyllables produced in isolation and non-isolation contexts

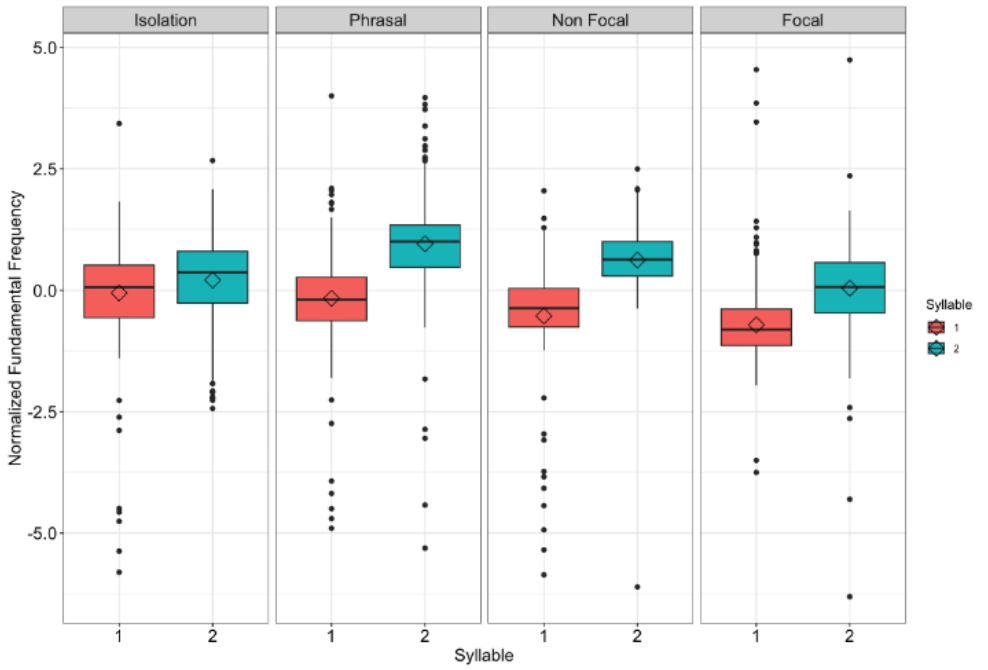

Figure 9: Distribution of mean *f*0 in Mundari disyllables produced in isolation and non-isolation contexts

Trisyllabic and tetrasyllabic words also show final lengthening in words said in isolation but there is no consistent pattern for vowel intensity or duration in non-isolation contexts (Figure 10-13).

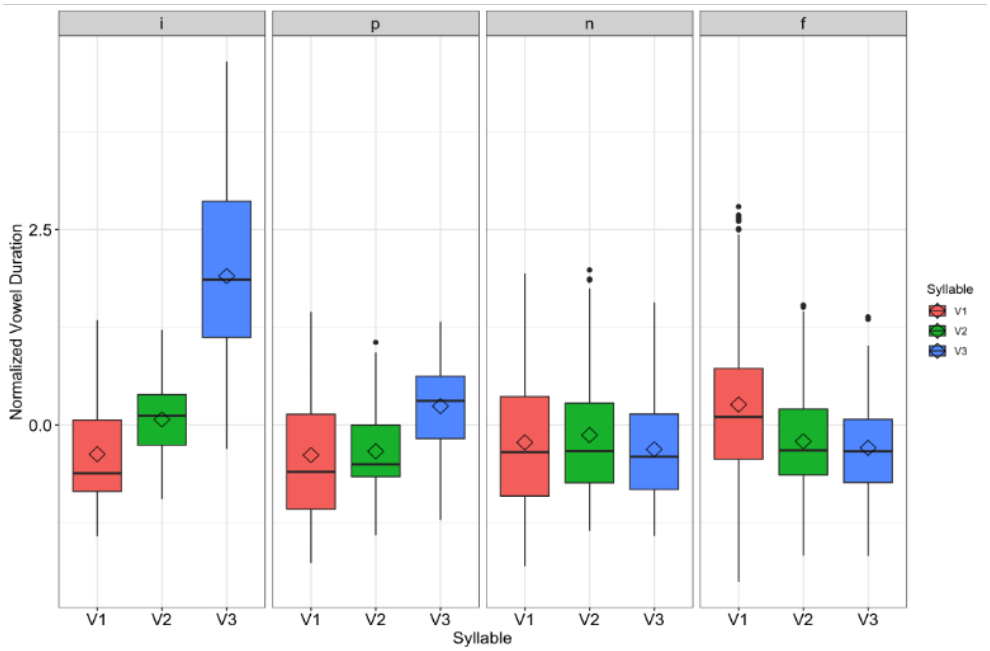

Figure 10: Distribution of vowel duration in Mundari trisyllables produced in isolation, phrasal, non-focal and focal contexts

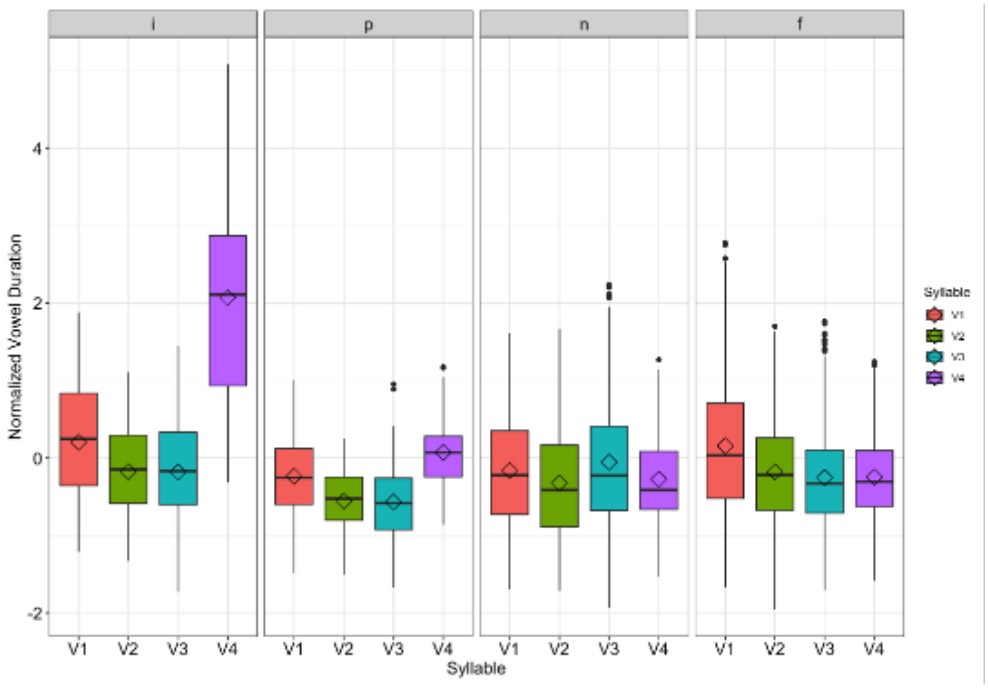

Figure 11: Distribution of vowel duration in Mundari tetrasyllables produced in isolation, phrasal, non-focal and focal contexts

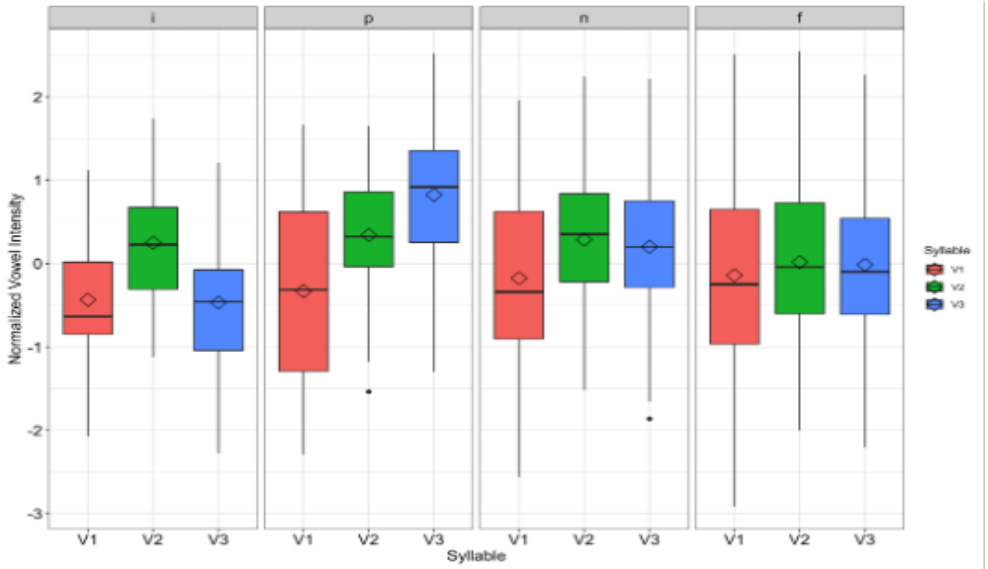

Figure 12: Distribution of vowel intensity in Mundari trisyllables produced in isolation, phrasal, non-focal and focal contexts

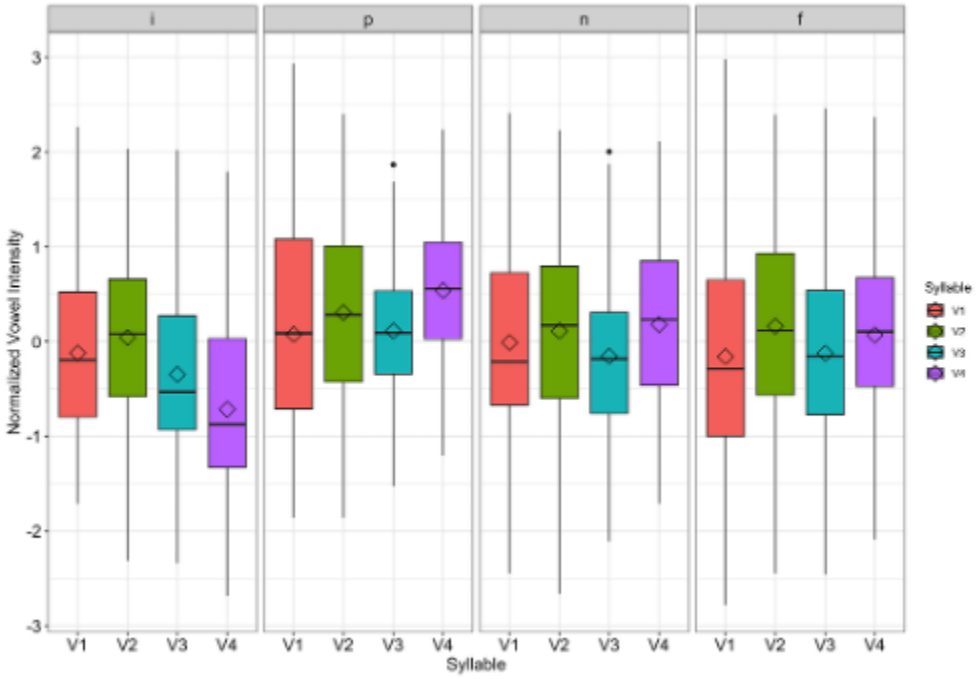

Figure 13: Distribution of vowel intensity in Mundari tetrasyllables produced in isolation, phrasal, non-focal and focal contexts

Moreover, as Figures 14-17 show, there is a confounding lack of patterning for five and six syllable words in Mundari with respect to either acoustic cue.

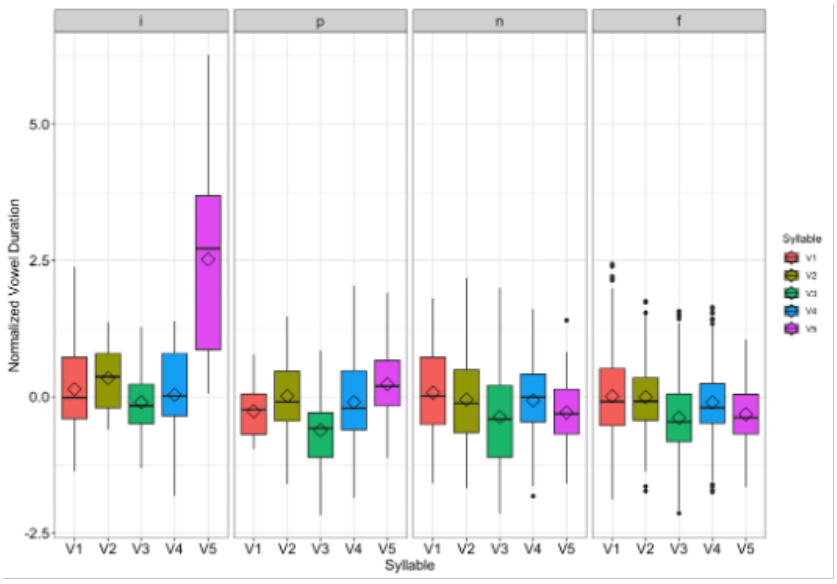

Figure 14: Distribution of vowel duration in Mundari pentasyllables produced in isolation, phrasal, non-focal and focal contexts

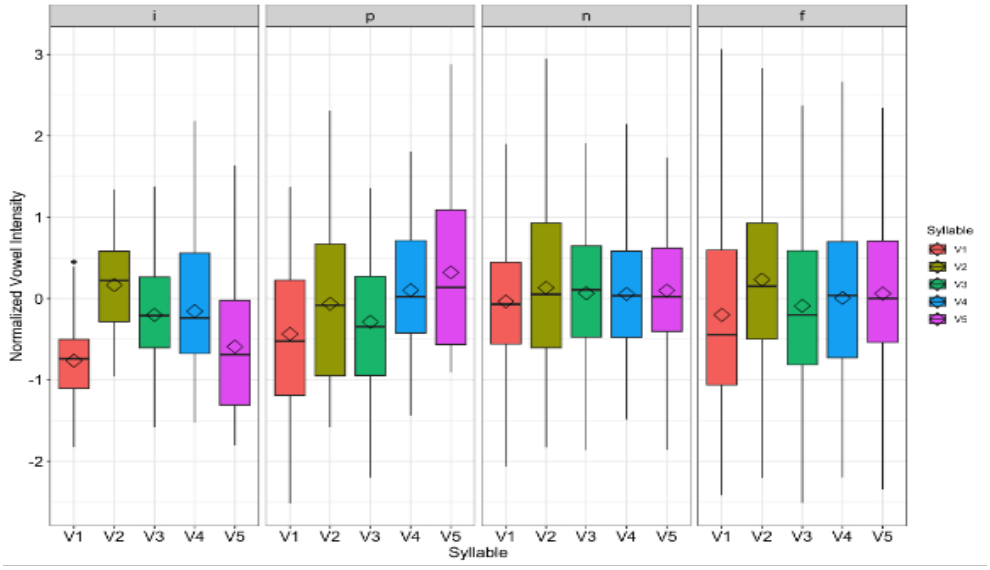

Figure 15: Distribution of vowel intensity in Mundari pentasyllables produced in isolation, phrasal, non-focal and focal contexts

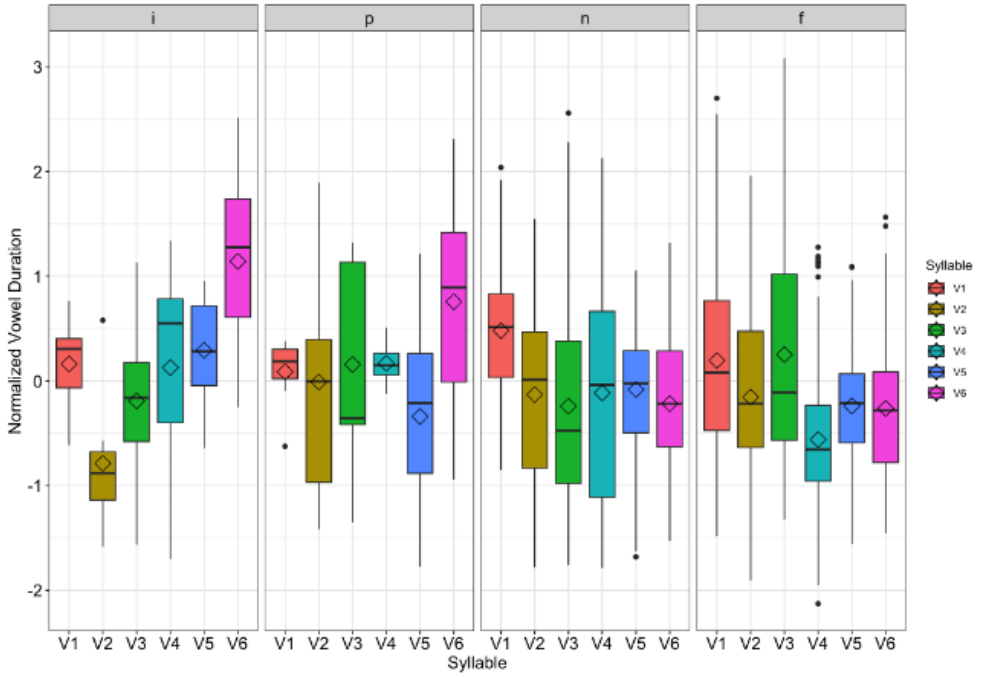

Figure 16: Distribution of vowel duration in Mundari hexasyllables produced in isolation, phrasal, non-focal and focal contexts

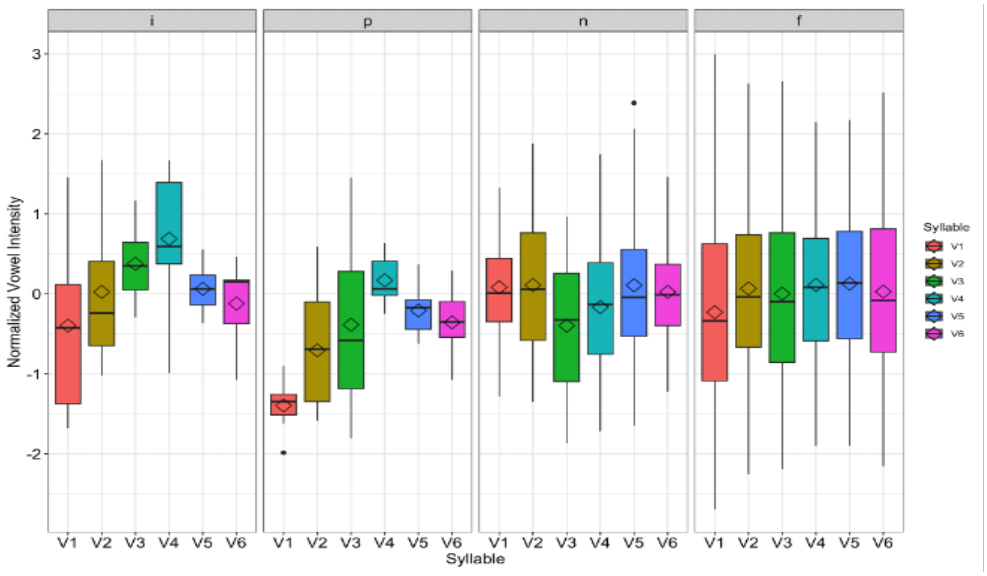

Figure 17: Distribution of vowel intensity in Mundari hexasyllables produced in isolation, phrasal, non-focal and focal contexts

Pitch, cued by fundamental frequency, shows nothing better either, whether trisyllables, tetrasyllables, pentasyllables or hexasyllables (Figures 18-21). These facts and others led us to propose that there is no level of phonological word in the phonoprosodic architecture of Mundari (Gogoi et al., 2024; Horo et al., 2024/2025). Note that while phonological word is not a category per se in Mundari, foot is the domain for a robust harmonic restriction against high and mid vowels. This domain is not, however, necessarily coterminous with the word or stem in Mundari.

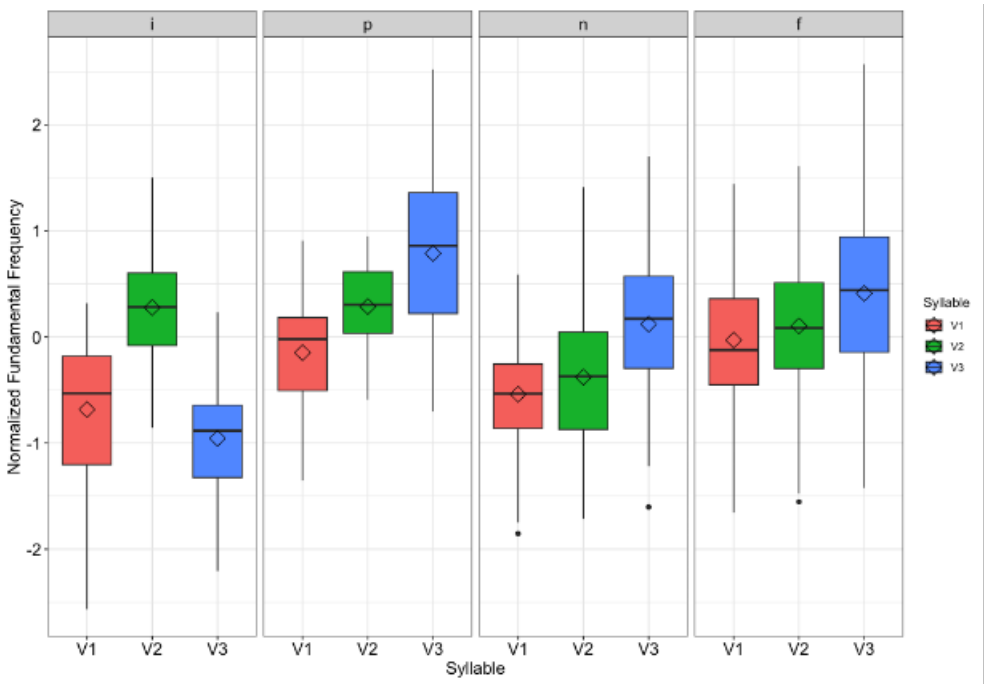

Figure 18: Distribution of mean F0 in Mundari trisyllables produced in isolation, phrasal, non-focal and focal contexts

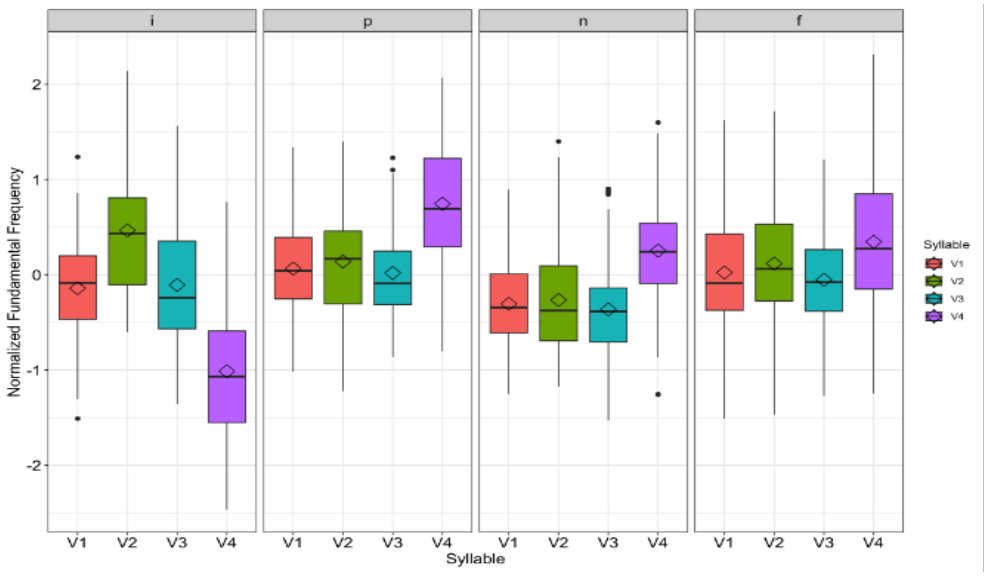

Figure 19: Distribution of mean F0 in Mundari tetrasyllables produced in isolation, phrasal, non-focal and focal contexts

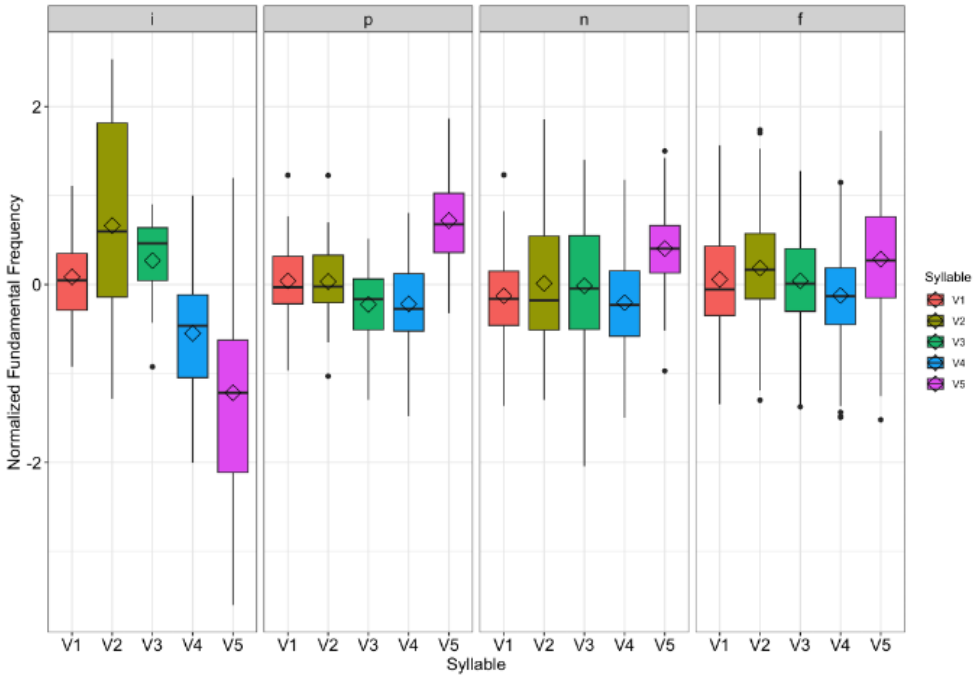

Figure 20: Distribution of mean F0 in Mundari pentasyllables produced in isolation, phrasal, non-focal and focal contexts

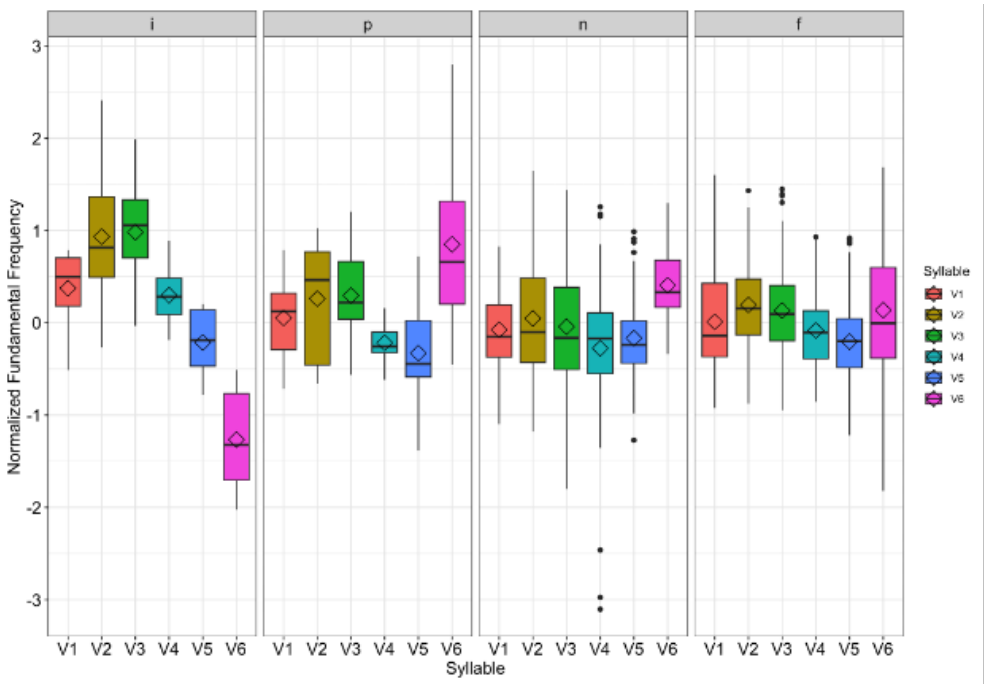

Figure 21: Distribution of mean F0 in Mundari hexasyllables produced in isolation, phrasal, non-focal and focal contexts

Put differently, there are no phonoprosodic processes or patterns in acoustic cuing that can justify any level larger than foot and smaller than utterance in Mundari, and none of the usual cues seem to have any role in information structuring at all. Perhaps this is predictable due to the presence and use of both morpholexical and syntactic means to encode information structure categories like topic or focus that are robustly used in Mundari. To be sure, there are general tendencies in utterances to see an increase in duration in final syllables, and a slight decrease in intensity and a decrease in pitch as a rule.

5 Summary

Munda languages may use morphological and syntactic means of encoding focus, but prosodic means are weakly developed and typically must be accompanied by morpholexical indices of information structure. They are each logically independent but can be interconnected. Individual languages show some commonalities, e.g., syntactically SOV marks subject focus and OSV object focus in both Sora and Mundari. But there are also differences found in the information structuring properties of scrambled syntactic configurations too, e.g., verb-initial order signals focusing on the verb or verb phrase in Sora, but VSO rather focuses on the subject and VOS on the object in Mundari. Some morphological markers might be old in the family in IS functions too, e.g., *do*, *ge*, *-de/i(ʔ)*, possibly others. Prosodic means of information focus in Sora interrogatives may show pitch peaks on the requested new information. Intensity is the cue to lexical prominence and generally peaks on the second syllable. Mundari has no prosodic cuing of prominence at the lexical level at all. Thus, Sora and Mundari prefer to utilize scrambled syntax and/or morphological indices encoding information structure together or independently rather than combining such with prosodic cuing of focalization for example, as commonly occurs with cleft constructions or scrambled syntax in English focalization.

Moving forward we will continue to map prosodic structures onto words (i.e., p- and g-words), phrases and sentences, as well as investigate the role of prosody in encoding information structure relations like focus in texts collected in four more Munda languages Kharia, Remo, Juang and Gtaʔ. This spans the full genetic spectrum within Munda. Once completed we will have a solid understanding of the formal and functional properties of information structuring in this important family of Austroasiatic languages of South Asia. The present study is just a first step into understanding the ways that Munda languages do and do not structure information. Once complete, we will have a clear idea the varied formal systems of prominence and focus attested in the family are and begin to be able to piece together how the individual attested systems may have arisen and changed over time.

Acknowledgment

Thanks to the US National Science Foundation grants #2041248 (Santali, Mundari, Kharia, Juang, Remo, Gorum, Gtaʔ) #1844532 (Sora) & the US National Endowment for the Humanities #PD-281083-21 (Santali, Mundari, Kharia, Juang, Remo, Gorum, Gtaʔ)

References

Asher, Nicholas. 2004. Discourse topic. In *Theoretical Linguistics 30*, 163–201.

Bearth, Thomas. 1999. The contribution of African linguistics towards a general theory of focus: Update and critical review. *Journal of African Languages and Linguistics* 20. 121—156. https://doi.org/10.1515/jall.1999.20.2.121.

Büring, Daniel. 1997. *The Meaning of Topic and Focus: The 59th Street Bridge Accent*. London: Routledge.

Büring, Daniel. 2010. Toward a typology of focus realization. In Malte Zimmermann & Caroline Féry (eds.), *Information Structure*, 177—205. Oxford: Oxford University Press.

Caron, Bernard. 2000. *Topicalisation et focalisation dans les langues africaines*. Louvain: Peeters [Afrique et Langage 1].

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles Li (ed.), *Subject and Topic*, 25—55. New York: Academic Press.

Dik, Simon C. 1980. On the typology of focus phenomena. *GLOT: Leids Taalkundig Bulletin* 3. 21—40.

Dik, Simon C. 1981. *Functional Grammar*. Amsterdam/New York: Oxford: North-Holland Publishing Company [North-Holland Linguistic Series 37].

Dik, Simon C. 1997. *The Theory of Functional Grammar I: The Structure of the Clause*. Berlin: Mouton de Gruyter.

Dik, Simon C, Maria Hoffmann, Jan R. de Jong, Djiang Sie Ing, Harry Stroomer & Lourens de Vries. 1981. On the typology of focus phenomena. In Harry van der Hulst Teun Hoekstra & Michael Moortgat (eds.), *Perspectives on Functional Grammar*, 41—74. Dordrecht: Foris Publications.

Drubig, Hans Bernhard & Wolfram Schaffar. 2001. Focus Constructions. In Martin Haspelmath et al (ed.), *Language Typology and Language Universals*, 1079—1104. Berlin: Mouton de Gruyter [Handbücher zur Sprach- und Kommunikationswissenschaft 20.2].

Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information Structure: Theoretical, Typological, and Experimental perspectives*, 234—257. Oxford, New York: Oxford University Press. https://doi.org/10.18452/13433.

Gogoi, Pamir, Luke Horo & Gregory D. S. Anderson. 2024. Phonetic and Phonological Analysis of the Mundari Vowel System. In P. Sidwell (ed.), *Austroasiatic linguistics. in honour of gérard diffloth (1939-2023)*, 75–90. Chiang Mai: Myanmar Center, Chiang Mai University.

Gundel, Jeanette & Thorstein Fretheim. 2004. Topic and focus. In Larry Horn & Gregory Ward (eds.), *The Handbook of Pragmatics*, 175—196. Malden, MA: Blackwell Publishers.

Horo, Luke & Gregory D. S. Anderson. 2021a. Prosody and Morphosyntax in Sora: A preliminary study. In *Proceedings of 1st International Conference on Tone and Intonation (tai)*, 51–55. doi:10.21437/TAI.2021-11.

Horo, Luke & Gregory D. S. Anderson. 2021b. Towards a prosodic typology of Kherwarian Munda languages: Santali of Assam. In *Papers from the 30th Annual Meeting of the Southeast Asian Linguistics Society*, vol. 8, 298–317. JSEALS Special Publications.

Horo, Luke & Gregory D. S. Anderson. 2023a. A preliminary study on prominence in Assam Santali disyllables. In Gregory D. S. Anderson (ed.), *Munda Linguistics. Typological, Descriptive and Diachronic Perspectives*, 127–150. Pune: Deccan College.

Horo, Luke & Gregory D. S. Anderson. 2023b. Towards a prosodic typology of Kherwarian Munda languages: Hasadaʔ and Naguri Mundari. Presented at Northeast Indian Linguistics Society. (NEILS).

Horo, Luke, Gregory D. S. Anderson & Pamir Gogoi. 2024/2025. Can Vowel Height and Onset Duration Cue Prominence in Mundari? In *Papers from the 33rd Annual Meeting of the Southeast Asian Linguistics Society*, vol. 13, 137—151. JSEALS Special Publications.

Horo, Luke, Gregory D. S. Anderson, Aman Singha, Ria Borah Sonowal & Opino Gomango. 2023a. Acoustic phonetic properties of p-words and g-words in Sora. In *Proceedings of (formal) Approaches to South Asian Linguistics*, vol. 12, https://ojs.ub.uni-konstanz.de/jsal/index.php/fasal/article/view/290.

Horo, Luke, Pamir Gogoi & Gregory D. S. Anderson. 2023b. Prominence in Mundari in disyllables and inflected polysyllables. Presented at HisPhonCog.

Horo, Luke, Pamir Gogoi, Aman Kumar Singha, Ria Borah Sonowal, Kelsey Bialo, Bikram Jora & Gregory D. S. Anderson. 2023c. Prominence in Mundari in disyllables and inflected polysyllabic nouns. Presented at SEALS.

Horo, Luke, Priyankoo Sarmah & Gregory D. S. Anderson. 2020. Acoustic Phonetic Study of the Sora Vowel System. *Journal of the Acoustical Society of America* 147. 3000–3011. https://doi.org/10.1121/10.0001011.

Horo, Luke, Aman Kumar Singha, Ria Borah Sonowal & Gregory D. S. Anderson. 2022. Acoustic phonetic properties of polysyllabic Sora words. Presented at ICoLSI.

Jacobs, Joachim. 2001. The dimensions of topic-comment. *Linguistics* 39. 641–681.

Krifka, Manfred. 2004. The semantics of Questions and the Focusation of Answers. In Gordon C. Lee & Daniel Büring (eds.), *Topic and Focus: A Cross-Linguistic Perspective*, 139–151. Dordrecht: Kluwer Academic Publishers.

Krifka, Manfred. 2007. Basic Notions of Information Structure. In Caroline Féry et. al (ed.), *Interdisciplinary Studies on Information Structure (isis) 6*, 13–56. Potsdam: Universitäts-Verlag [Working papers of the SFB 632].

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: A theory of topic, focus, and the mental representations of discourse referents*. New York, NY, USA: Cambridge University Press.

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

Sasse, H.-J. 1987. The thetic-categorical distinction revisited. *Linguistics* 25. 511—580.

Uhmann, Susanne. 1991. *Fokusphonologie*. Tübingen: Niemeyer.

van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in Language* 40. 259—301.

Zimmermann, Malte. 2008. Contrastive focus and emphasis. *Acta Linguistica Hungarica* 55. 347—360.
Zimmermann, Malte. 2016. Predicate focus. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford: Oxford University Press.
Zimmermann, Malte & Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121. 1651—1670.
Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. The MIT Press, MA: Cambridge dissertation.