

Information structure marked by morphology, syntax and prosody in Munda languages

Gregory D. S. Anderson, Bikram Jora and Luke Horo
Living Tongues Institute for Endangered Languages
livingtongues@gmail.com

Munda languages represent the westernmost branch of the Austroasiatic language phylum spoken primarily in the Indian states of Jharkhand and Odisha but also found in significant numbers in Madhya Pradesh, Maharashtra, Chhatisgarh and West Bengal, as well as largely in diaspora in Assam and a few other states, in addition to Nepal, Bhutan and Bangladesh. Typologically Munda languages are quite distinct from the majority of Austroasiatic languages as they are syntactically verb-final and morphologically agglutinative or polysynthetic, with many polysyllabic words, and representing a subtype of peripheral South Asian areal complex (shared for example with Kiranti of the Trans-Himalayan (Tibeto-Burman) phylum and the isolate language Burushaski). Most other Austroasiatic languages on the other hand constitute core members of the Mainland Southeast Asian areal group, with sesquisyllabic or even monosyllabic words, and show verb-medial (and occasionally verb-initial) structure syntactically and make limited use of derivational morphology and extremely rarely use inflectional morphology at all (albeit some do in fact make limited use of such features). Munda languages make use of a range of formal devices to encode focus. This includes morphological elements (1)-(2), e.g., in Mundari, or syntactic processes such as scrambling (3), as in Sora (which also has non-cognate morphological contrastive topic and focus markers).

(1) *-ge* focalizer in Mundari

i. *aŋ-ge=ŋ senog-a*

I-FOC=1SG go-IND

‘(It is) I (who) will go’ (Osada 1992: 136-7)

ii. *mandi-ge=ŋ dʒom-ke-n-a*

food-FOC=1SG eat-AOR-I/M-IND

‘I ate the food’

iii. *Ranci-ate-ge=m hidʒuʔ-aka-n-a*

Ranchi-ABL-FOC=2SG come-ANT-I/M-IND

‘you have come from Ranchi’

iv. *mandi=ko dʒom-tan-ge-a*

food=3PL eat-IPFV-FOC-IND

‘they are eating food’

(2) Santali *-do* and *-ge*

i. *nui gəi-dɔ-e baŋ kan-t-iŋ-a iŋ-ren-dɔ-e uni uni kan-t-iŋ-a*
this cow-TOP-3SG NEG COP-POSS-1SG-IND I-GEN-TOP-3SG that RDP COP-POSS-1SG-IND

‘This cow is not mine; mine is that one there.’ (Neukom, 2001:34)

ii. *iŋ-reak-dɔ alo-ge-pe bɔtɔr-ok’-a*

I-GEN-TOP PROH-FOC-2PL fear-MID-IND

‘Do not be afraid of me at all.’ (Neukom, 2001:30)

iii. *banar-ge tuŋi-kin-me*

both-FOC pierce-3DL:OBJ-2SG:IMP

‘Pierce them both.’ (Ghosh, 2008:47)

iv. *ʃalaʔ-ge-a-ŋ*

GO:IPFV-FOC-IND-1SG

‘I will certainly go.’ (Campbell 1899-1902:184)

(3) Sora

i. *opino-n kinte-n*

dʒum-le

ii. *kinte-n opino-n dʒum-le*

banana-N.SFX Opino-N.SFX eat-PST
'Opino ate the banana'

Opino-N.SFX banana-N.SFX eat-PST
'Opino ate *the banana* (not the mango)'

iii. *opino-n dʒum-le kinte-n*
Opino-N.SFX eat-PST banana-N.SFX
'Opino ate the banana (not Indamo)'

iv. *dʒum-le opino-n kinte-n*
eat-PST Opino-N.SFX banana-N.SFX
'Opino ate the banana (he did not grow/pick it)'

In this paper we examine the systems of focus marking in three languages of the Munda family, viz., Mundari of the Kherwarian branch and its closely related sister language Santali (Neukom 2001) and the distantly related Sora of the Sora-Gorum subgroup of Munda (Sidwell 2015). Mundari and Santali comparisons are used to show low-level variation across closely related lects, where the elements *-do* and *-ge* seem to overlap in function in certain contrastive contexts, with *-ge* more promiscuous in its host selection appearing with both arguments and predicates while *-do* is rather found more commonly on arguments. Both are found in the distantly related Sora with different synchronic functions (4), but likely with earlier roles in encoding information structure. In this paper we describe the systems of focus in these languages individually and compare the three. We also observe what role intonational properties can be manipulated in these languages to encode information structure as well and thus detail prosodic prominence signaled by various acoustic cues when possible. Where relevant, we draw attention to commonalities with or divergences from other Austroasiatic languages. All data come from the authors' field notes and findings are substantiated with instrumental analysis of speech data.

(4) i. Tengkalasor

<i>kən=do</i>	<i>aboi</i>	<i>(e)le</i>	<i>p^haʔto</i>	<i>naŋ-li(n)=den=dem</i>
this=DISC	one	may	study	learn/get-PST-N.SFX=COND=COND

<i>baŋsa=biŋ</i>	<i>de-to</i>	<i>tiki</i>	<i>dʒana</i>
good=MOD	MOD-NPST	later	also

'If one may learn by studying, that could be good later too'

ii. West Bengal Sabar

<i>təb=na</i>	<i>a-undʒi-n-len</i>	<i>daku-li kudubə-mər-an</i>	<i>bagu</i>
then=FOC	DEP-four-N.SFX-1PL	COP-PST total-people-N.SFX	two

<i>nen</i>	<i>dangari</i>	<i>oʔon-nen</i>	<i>dadigti=gi</i>
1SG	youth.FEM	child-1SG	that's.all=TOP

'we were surely then four people in total, only two of us four (remain now), me and my daughter, that's all (of us now)'

References:

- Campbell, A. 1899-1902. A Santali-English Dictionary. Manbhum: The Santal Mission Press.
Ghosh, A., 2008. 'Santali', in G. D. S. Anderson (ed.), *The Munda Languages*, New York: Routledge, pp. 11–98.
Neukom, Lukas. 2001. *Santali*. München: LINCOM.

Osada, T. 1992. *A Reference Grammar of Mundari*. Tokyo: TUFS.

Sidwell, Paul. 2015. Austroasiatic Classification. In M. Jenny and P. Sidwell (eds.) *The Handbook of Austroasiatic Languages*. Brill: Leiden/Boston, pp. 144-220.