# S Berkeley Linguistics Society

South-East Asian Features in the Munda Languages: Evidence for the Analytic-to-Synthetic Drift of Munda Author(s): Patricia Jane Donegan and David Stampe Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics (2002), pp. 111-120

Please see "How to cite" in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <a href="http://linguistics.berkeley.edu/bls/">http://linguistics.berkeley.edu/bls/</a>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via <u>eLanguage</u>, the Linguistic Society of America's digital publishing platform.

# South-East Asian Features in the Munda Languages: Evidence for the Analytic-to-Synthetic Drift of Munda

# PATRICIA JANE DONEGAN AND DAVID STAMPE

Linguistics, University of Hawai'i at Mānoa

This paper, written in memory of Eugénie J. A. Henderson and A. K. Ramanujan and read at the celebration of James Matisoff on his retirement, expresses our thanks for all their deep insights into the languages of South and South-East Asia.

# 1. Opposite Orders of Thought

The Munda (South Asian) and Mon-Khmer (South-East Asian) branches of the Austroasiatic language family are so exactly opposite at every level of structure that Sir George Grierson in his *Linguistic Survey of India* remarked that if they were descended from a common language, the language must have been adopted by peoples with opposite orders of thought (1904: v. 2, p. 2).

In (1) is a listing of typological oppositions between Munda and Mon-Khmer, adapted from Donegan & Stampe 1983. That paper showed how their opposite synthetic vs. analytic traits might be explained as due to polar drifts driven by their opposite – falling vs. rising – phrase and word rhythms.

(1)	MUNDA	MON-KHMER		
Phrase Accent:	Falling (initial)	Rising (final)		
Word Order:	Variable – OV, AN, Postpositional	itional Rigid – VO, NA, Prepositional		
Syntax:	Synthetic – subj/obj agreement on verb	Analytic – no inflectional morphology		
Word Canon:	Trochaic Iambic, monosyllabic			
Morphology:	Agglutinative, Suffixing, Polysynthetic Fusional, Prefixing or Isolating			
Timing:	Isosyllabic or isomoraic	Isoaccentual		
Syllable Canon:	(C)V(C)	Unaccented (C) $\theta$ , accented (C)(C)V(G)(C)		
Consonantism:	Stable, Geminate clusters	Shifting, Tonogenetic, Non-geminate		
		clusters		
Tone/Register:	Level tone (Korku only)	Contour tones or registers		
Vocalism:	Stable, monophthongal, harmonic	Shifting, diphthongal, reductive		

BLS 28S (2002): 111-120

We will review the polarizing effects of falling vs. rising accent in section 2.

In sections 3 and following, we discuss similarities of Munda to Mon-Khmer, and argue that these must be retentions or developments from an originally rising typology, and therefore that proto-Austroasiatic was of the rising type – that it was analytic like Mon-Khmer, not synthetic like Munda.

Some linguists view a spontaneous shift of type from analytic to synthetic as impossible, and hold that a holistic drift, as from early to modern Indo-European, must arise in the phonetic decay of suffixes, with a change from synthetic to analytic grammar, and an accompanying drift from OV to VO word order, etc. On their view, the synthetic structure of Munda would have to be reconstructed for proto-Austroasiatic and then lost in Mon-Khmer. They might even argue that the loss was due to the areal influence of the analytic languages of SE Asia.

Or they might hold that a change from analytic to synthetic can occur only under the influence of synthetic languages. It has often been asserted, e.g. in the 1978 *Encyclopedia Britannica* article on Austroasiatic languages, that Munda synthetic structure must be due to the influence of the synthetic languages of South Asia. But Indo-Aryan and Dravidian are modifier-marking, in Nichols' terminology (1992), while Munda is head-marking, and even if one does not accept this dichotomy as immutable, it is hardly likely that modifier-marking languages could induce analytic languages to become head-marking languages.

# 2. Polar rhythms and polar drifts

The main reason the divergent structures of Munda and Mon-Khmer cannot be explained as due to convergence with other languages in their respective South or SE Asian language areas is that Munda and Mon-Khmer, and other South and SE Asian languages, do not just differ in structure: they are *opposite* at every level of structure. Such a polarization can be explained only by a linguistic principle, not a historical one, and the fact that it pervades every level of structure, from lexicon to syntax to phonetics, points to the single opposition that pervades every level of language: the opposition of falling vs. rising rhythm.

Munda and Mon-Khmer accentuation are opposed in just this way. Munda languages have falling (initial) accent in phrases and in words, while Mon-Khmer has rising (final) accent in phrases and in words. In this section we will sketch our 1983 hypotheses about how this opposition guides the syntactic, morphological, and phonological drift.

Heads of phrases, as presupposed information, tend to be accentually backgrounded relative to modifiers, and so in consistent head-last languages, phrase accent is falling (initial), as in South Asia, while in consistent head-first languages, it is rising (final), as in South-East Asia. Perhaps head-last (left-branching) order poses problems for short-term memory, because falling languages augment word order with incorporation, as in Munda and Tibeto-Burman, or case marking, as in Indo-Aryan or Dravidian.

Word accent tends to adopt the falling or rising structures of phrases, falling (accented at or near the beginning of word) as in Munda and Dravidian, vs. rising (accented at or near the end of the word) as in Mon-Khmer and Tai. Grammatical elements are backgrounded relative to lexical elements, so they are treated as extrametrical or are affixed away from the accent, so some languages with initial accent like Dravidian or Finnic have only suffixes, and some languages with final accent like Mon-Khmer have only prefixes. VC- prefixes may be infixed before C-initial roots to avoid creating heavy syllables that would invite accent.

Rising accent gives an "iambic" word, really an anacrustic syllable plus a stressed syllable, allowing word- (stress-) timing; the initial vowel is reduced or omitted, forming monosyllables with initial clusters that invite consonant shifts and registers or contour tones on the bimoraic and highly diphthongizable final vowel (Matisoff 1973), e.g. Mon-Khmer \*[bəˈlu:] Khmer [ˈpləu] 'thigh'. Falling accent gives a "trochaic" word, both syllables within the bimoraic beat, inviting harmony (Munda [ˈbulu] 'id.') or apocope (bimoraic [ˈbul]), but as suffixes are piled on, isochrony at the word level becomes impossible, timing focuses on the syllable or mora, and vowels and consonants are far more stable than under stress-timing (Donegan 1993).

A holistic reversal of typology seems to require a reversal of accentuation. Germanic, Italic, and Celtic, for example, originally had head-last phrases, with falling accent, as is evident in the front-rhymed (alliterative) forms of their early verse, but they shifted to head-first phrase structure, with rising accent, and end-rhymed verse. Morphology lags behind: the ordering of compounds and affixes remains head-last in English long after phrases became head-first, and it might even be argued that the order of compounds like *blackbird* is what has retarded the reversal of adjective-noun phrases like *black bird*. But this lag can preserve a hint of the history, or prehistory, of a language.

The reversal of typology in Austroasiatic has been even more profound than in Indo-European: Munda languages are more synthetic than proto-Indo-European, and Mon-Khmer languages are far more analytic even than English. Perhaps this reflects a greater time-depth for Austroasiatic than for Indo-European.

Now we will proceed to the evidence that the reversal in Austroasiatic was opposite that in Indo-European – that proto-Austroasiatic had an analytic and head-first structure like that of Mon-Khmer, but that Munda drifted to a synthetic and head-last structure due to a reversal from rising to falling accent.

### 3. Vocabulary

**3.1.** Cognates. The evidence of the original linguistic unity of Munda and Mon-Khmer has rested, and still rests, mainly on lexical cognates. Though the vocabulary that we can reconstruct as Austroasiatic is far smaller than that for proto-Munda and proto-Mon-Khmer, and those are a magnitude smaller than the shared vocabularies of Indo-Aryan or of Dravidian, the Austroasiatic vocabulary is still solid enough to leave no doubt of the unity of the family. Despite losses

due to borrowing, we have Austroasiatic cognates for the basic verbs and nouns relating to body, family, home, field, and forest, and for pronouns, demonstratives, and numerals. Agricultural vocabulary points to a very early SE Asian homeland (Zide & Zide 1976), but that does not prove that proto-Austroasiatic was of the analytic type now identified as "South-East Asian".

#### 3.2. Word Structure

We compared Mon-Khmer "iambic" or monosyllabic words and Munda "trochaic" words in section 2. Evidence that Munda trochaic words derive from proto-Austroasiatic originals of the Mon-Khmer iambic type include (a) The tendency for  $C_1V_1C_2V_2(C_3)$  cognates of the type Mon-Khmer [bəˈluː] : Munda [ˈbulu] 'thigh' to show a harmonic  $V_1$  in Munda for the unaccented neutral  $V_1$  of Mon-Khmer, and (b) the high frequency of simple CVCV(C) words in Munda as opposed to equally admissible CVCCV(C) words.

#### 3.3. Affixation

Mon-Khmer and other Mainland SE Asian language families have only prefixes and infixes. This is peculiar to head-first languages, just as having only suffixes is peculiar to head-last languages like Dravidian and Finnic. Munda has prefixes and infixes cognate to Mon-Khmer, but it also has even more suffixes. If these suffixes had existed in proto-Austroasiatic, and had been lost in Mon-Khmer, we would not expect to find Mon-Khmer cognates for the Munda suffixes. In fact, we do find Mon-Khmer cognates, but they are independent words in Mon-Khmer. For example, Munda languages mark the plural of nouns and 3rd plural of verbs with suffixes like -ku, -ki, -gi, -ji. Mon-Khmer languages lack number suffixes, but many have free-standing 3rd plural pronouns like Khasi ki 'they'.

Munda languages mark possessive and object persons with suffixes, e.g. Sora /si?iŋ-lɛn/ 'our house (lit. house-us)', /ərgəl-da-tə-lɛn/ 'we're thirsty (lit. thirst-affect-nonpast-us)'. Mon-Khmer languages lack person suffixes, but they have free-standing personal pronouns cognate to the Munda suffixes. Here are examples from Pinnow's extensive 1965 study:

(3.3)	Proto-Munda	Mon-Khmer
'1 sg.'	*-iŋ	Pear, Bahnar in; Mon ai (oa); Srê ?an; Khmer ən
'2 sg.'	*-me	Khmer me, Bahnar mih, Srê mi, Khmu' mee
'1 pl.'	*-le/-ne	(See Pinnow 1965: 23ff.)
'2 pl.'	*-pe	Palaung pε, Riang pe?, Mon beh (pih), Wa pu:-i,
		Khasi phi

The change of free pronouns to clitics and affixes is commonplace, but the change of affixes or clitics to free forms is not. We conclude that Munda suffixes derive from the synthesis of independent words, as in Mon-Khmer.

# 3.4. Compounding

In relatively recent compounds, Munda structure is head-last, e.g. Sora /bəŋsa-im/ 'good chicken', but in older compounds, the structure is head-first, e.g. Sora /əd̞rɛ-im/'chicken egg (lit. egg-chicken)', /gəd-im-na/ 'sacrificing a chicken (lit. cut-chicken-nom.)'. The latter is clearly a reflex of Mon-Khmer-type head-first compounding.

The accentuation of compounds illustrates the natural principle that heads are accentually subordinate to modifiers. Thus the head-first compounds of Mon-Khmer show the characteristic rising rhythm of Mon-Khmer words and phrases, while the head-last compounds of Munda show the falling rhythm of Munda words and phrases, and indeed in some languages like Sora the second element of the compound is synchronically limited to one syllable.

That falling rhythmic patterns have been imposed on older Munda compounds with head-first structure, like /ˈəsəŋ - ˌim/ (lit. *feces chicken*, 'chicken dung' – compare the full form of *chicken*, /ˈkənsim/), shows that this word order existed before Munda adopted the falling rhythm typical of the South Asian area and of head-last languages generally.

#### 4. Phrasing

#### 4.1. Syntax

Munda phrase structure is consistently head-last, with SOV and AN order, and postpositions. Mon-Khmer phrase structure is just as consistently head-first, with SVO (rarely VSO) and NA order, and prepositions.

(4.1a) 'Monosi/Saran went to the market; he bought rice.'

```
Sora: Monosi bajar -ban yer -r -ε; ənin roŋko-n ɲi -l -ε
Monosi market-to go -past -3sg; he rice -art buy -past -3sg.
Khmer: saraan tɨw dal psaa; kŏət tɨŋ ʔaŋkaa
Saran go to market; he buy rice
```

```
(4.1b) 'I don't want to eat all the fish.'
```

```
Sora: nɛnkuddɨb əyə -n -ədəŋ ə- jəm -ben idsɨm -t -ay ted
I all fish -art-obj inf-eat -inf want -pres -lsg not
Khmer: khnom mɨn caŋ nam trəy cap
```

Khmer: khnom min can nam trəy cap I not want eat fish all

# 4.2. Polysynthetic Morphology

Words are more resistant to internal changes of accent and ordering than phrases. We have already noted that noun compounds in Munda retain a head-first order. The Munda verb, which is polysynthetic, likewise shows internal head-first order, as if head-first phrases of the Mon-Khmer type were fused, with no order change.

```
(4.2a) 'He didn't give me rice':
```

```
(3past)
        he
               not
                    give
                           rice
                                 me
                           dar-
Sora:
        anin
               əd-
                    tiy-
                                 in-
                                        ten
Khmer: kŏət
                                 khnom
               min aoy
                           baay
Sre:
        khay
               ?à?
                    ?ay
                          pyan
                                  ?an
```

Similarly for the sentence cited in (4.1b) in the Sora "syntactic" style, but in (4.2b) in its more idiomatic (and older) "morphological" (polysynthetic) style:

```
(4.2b) 'I don't want to eat all the fish.'
```

```
I not want eat fish all (-pres. -intr. -1p.)
Sora: nen əd- məl- jom -yə -aj -t -en -ay
Khmer: khnom min can nəm trəy cap
```

# 5. Phonology

#### 5.1. Vowels

The vowel systems of Mon-Khmer and SE Asia generally are among the most complex in the world, and even at the proto-Mon-Khmer level they share the peculiarity of having three central or back unrounded vowels (Shorto 1976). Munda vowel systems mostly appear as triangular systems of five vowels, like the typical systems of Indo-Aryan and Dravidian languages, but a striking exception is the Sora system, whose three central vowels look very Southeast Asian:

```
(5.1) high i i u e.g. [id-] 'scratch', [id-] 'fan', [-lud-] 'ear' mid tense e o [-əd-] 'thorn', [-lod-] 'cord' mid lax ε ə ɔ [εd-] 'roll', [əd-] 'prop', [ɔd-] 'knead' low a [ad-] 'drive (cattle)'
```

More striking is the fact that at the lowest levels of reconstruction, it is necessary to reconstruct three central vowels for every Munda subgroup: Sora-Gorum

(A. Zide 1982), Gutob-Remo (N. Zide 1965), Kharia-Juang (Stampe 1978), Kherwarian (Munda 1969), and Korku-Kherwarian (N. Zide 1965). These point to a proto-Munda and proto-Austroasiatic system like that of proto-Mon-Khmer.

Some South Munda languages have vowels with glottals, as in Sora, where  $[V_i ? V_i]$  in free forms alternates with  $[V_i]$  in combining forms, e.g.  $[ja?a\eta]$  'bone' beside [a-palaŋ-jaŋ] 'broken bone', and in 1965 Norman Zide proposed that these and a number of vowel and consonant puzzles in Munda history might be solved by a proto-Munda series of laryngealized vowels. In 1989 Diffloth gave evidence of creaky-voiced vowels in proto-Mon-Khmer. Vowel registers are rare in South Asia but common in SE Asia; if the correspondences can be resolved, this would be another Mon-Khmer-like feature of Munda.

#### 5.2. Consonants

Indo-Aryan languages have released final consonants, and Dravidian languages end words with an "enunciative" vowel. In contrast, Munda languages typically have unreleased final consonants. In older handbooks these were called implosive, in the sense of "not plosive" rather than "inwardly plosive", which led some phonological surveys to count them wrongly as ingressive; they are just unreleased, glottalized, and voiceless as in English *cat* [kæt'], Cockney [kæ?]. This "checking" of final stops is commonplace in Mon-Khmer and other mainland SE Asian languages. Presumably it was a proto-Austroasiatic feature, because while proto-Mon-Khmer and proto-Munda had voiced as well as voiceless stops nonfinally, there is no evidence of more than one series finally. In the absence of suffixes, as in Mon-Khmer, the invariably checked final stops are lexically voiceless. But in Munda, final stops before vocalic suffixes alternate with their voiced equivalents, as in these Sora examples:

Nonfinal voiceless and voiced stops contrast before vowels (e.g. [əpan] 'easily' vs. [əba-n] 'mohwa tree'), but as the phonemic notations above indicate, the final voiceless stops are identified not with the voiceless non-final stops but with the voiced ones. Even a voiceless final stop in a foreign word like English *pipe*, when suffixed with the article /-ən/, is revealed as voiced: Sora [pa.ɛ.bən].

In Mundari, a similar analysis of stops in Mundari causes some speakers to produce final stops as nasally released voiced stops, e.g. [dup'] ~ [dub<sup>m</sup>] 'to sit' (Osada 1992), and English *David* as [de.bit'] ~ [de.bid<sup>n</sup>] (personal observation).

In fact, one Munda language, Juang, has lost final checking, and its previously checked and voiceless final stops have emerged as voiced, not voiceless:

(5.2b)	Juang		Kharia		Sora	
'head'	/bokob/	[bo.kob]	/bokob/	[bo.kop']	/bokob/	[bə.?əp']
'mouth'	/tomod/	[to.mod]	/tomod/	[to.mot']	/tomod/	[tə.ʔəţ']
'water'	/dag/	[dag]	/dag/	[da?]	/dag/	[da.?a]
'die'	/goj/	[godz]	/goj/	[goi']	/goj/	[goi']

For the lexical representation of *all* the morpheme-final stops in a language to be perceived as voiced, even though they are voiceless except before a vowel, is quite extraordinary, but it is clearly a fact of Munda. As to how the prevocalic forms of the stops became voiced, we believe that it was by exactly the same process as in Sanskrit, where word-final stops were voiced before vowels. Proto-Austroasiatic, like Mon-Khmer, had both voiceless and voiced stops initially, but only voiceless stops finally, and the final stops were invariably voiceless because there were no affixes to block devoicing. But when Munda began to use clitics and suffixes, word-final stops must still have been syllable-final, i.e. VC#V = VC.V, and in syllable-final but intervocalic position they assimilated voicing, just as Sanskrit did in word-final (presumably syllable-final) stops in external sandhi:

Non-word-finally, intervocalic stops have the natural syllable division V.CV and, exactly as in Sanskrit (e.g. *pi.tah* 'father'), they do not become voiced:

What is significant for our thesis is that this voicing of stops before vocalic suffixes in Munda could only have occurred when Munda joined syllable-final checked stops to vocalic suffixes. As soon as the suffixes became integral parts of words, the syllabication of stops was naturalized to V.CV in all the Munda languages (see the examples in 5.2a–d above). So the reinterpretation of final stops as voiced must be a reflex of the moment when Munda languages crossed over from a non-suffixing Austroasiatic morphology (like that of Mon-Khmer) to a suffixing morphology.

#### 6. Summary

We have argued that proto-Austroasiatic had the same analytic, head-first, and rising structure as its daughter Mon-Khmer and other mainland South-East Asian languages, and that the Munda languages have preserved clear evidence of that structure even as they evolved toward the synthetic, head-last, falling structure typical of other South Asian languages. That does not necessarily mean that the speakers of proto-Austroasiatic were actually in South-East Asia, or that the Munda changes took place in South Asia. But it does mean that Munda is a clear example of a drift that was exactly the opposite of the drift that is familiar from Indo-European, toward analysis. Further, the drift of Munda was more complete than that of Indo-European, since it began with one of the most analytic structures among the languages of the world, and ended with one of the most synthetic.

# References<sup>1</sup>

- Diffloth, Gérard. 1989. Proto-Austroasiatic creaky voice. *Mon-Khmer Studies* 15: 139-154.
- Donegan, Patricia J. 1993. Rhythm and vocalic drift in Munda and Mon-Khmer. Linguistics of the Tibeto-Burman Area 16: 1-43.
- & David Stampe. 1983. Rhythm and the holistic organization of language structure. In John F. Richardson, Mitchell Marks, & Amy Chukerman (eds.) *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*: 337-353. Chicago: Chicago Linguistic Society.
- Grierson, George A. 1904. *Mon-Khmer and Siamese-Chinese families. Linguistic survey of India, Vol. 2.* Reprinted 1966, Delhi: Motilal Banarsidass.
- Henderson, Eugénie J. A. 1965. The topography of certain phonetic and morphological characteristics of South East Asian languages. In Milner & Henderson 2: 400-434.
- Huffman, Franklin E. 1986. *Bibliography and index of Mainland Southeast Asian languages and linguistics. (Yale Language Series.)* New Haven: Yale University Press.
- Jenner, Philip M., Laurence C. Thompson, & Stanley Starosta (eds.). 1976. *Austroasiatic studies*. (Proceedings of the First International Conference on Austroasiatic Linguistics.) 2 vols. Honolulu: University of Hawai'i Press.
- Masica, Colin. 1976. *Defining a linguistic area*. Chicago: University of Chicago Press.
- Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.), Consonant types and tone (Southern California Occasional Papers in Linguistics, 1): 71-95. Los Angeles: UCLA.

1

<sup>&</sup>lt;sup>1</sup> Uncredited Munda citations in this paper for languages not mentioned in the bibliography are from our own field notes and recordings. For areal and typological facts we have relied heavily on the indispensable works of Ramanujan and Masica for South Asia and Henderson for SE Asia.

- Milner, George B., & Eugénie J. A. Henderson (eds.). 1965. *Indo-Pacific linguistic studies*. 2 vols. Amsterdam: North-Holland.
- Munda, Ram Dayal. 1969. *Proto-Kherwarian sound system*. M. A. thesis, University of Chicago.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Pinnow, Heinz-Jürgen. 1959. Versuch einer historischen Lautlehre der Kharia-Sprache. Wiesbaden: Harassowitz.
- . 1960. Über den Ursprung der voneinander abweichenden Strukturen der Munda- und Khmer-Nikobar-Sprachen. *Indo-Iranian Journal* 4: 81-103.
- —. 1965. Personal pronouns in the Austroasiatic languages: a historical study. In Milner & Henderson 2: 3-42.
- Ramanujan, A. K., & Colin Masica. 1969. Toward a phonological typology of the Indian linguistic area. In T. A. Sebeok (ed.) *Current trends in linguistics* 5: 543-577. The Hague: Mouton.
- Shorto, Harry L. 1976. The vocalism of Proto-Mon-Khmer. In Jenner et al. 2: 1041-1067.
- Stampe, David. 1978. Proto-Central Munda phonology. Second International Conference on Austroasiatic Linguistics, Mysore, 1978.
- Zide, Arlene R. K. 1982. *A reconstruction of proto-Sora-Juray-Gorum phonology*. Ph.D. thesis, University of Chicago.
- \_\_\_\_ & Norman H. Zide. 1976. Proto-Munda cultural vocabulary: Evidence for early agriculture. In Jenner et al. 2: 1295-1334.
- Zide, Norman H. 1965. Gutob-Remo vocalism and glottalised vowels in Proto-Munda. In Milner & Henderson 1: 43-53.
- —, ed. 1966a. Studies in comparative Austroasiatic linguistics. The Hague: Mouton.
- —. 1966b. Korku low tone and the Proto-Korku-Kherwarian vowel system. In Zide 1966a 1: 214-229.

Linguistics Department University of Hawai`i at Mānoa 1890 East-West Road, Honolulu, HI 96822

donegan@hawaii.edu, stampe@hawaii.edu