

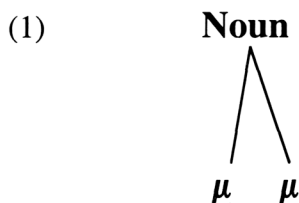
# Issues in Proto-Munda and Proto-Austroasiatic Nominal Derivation: The Bimoraic Constraint

Gregory D. S. Anderson and Norman H. Zide

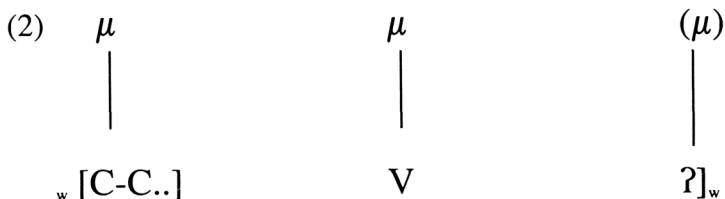
University of Manchester and University of Chicago

## 0 Introduction

When attempting to reconstruct proto-level forms of nouns in Munda, there is often a clear root-element which is generally monosyllabic and necessarily monomorphemic in form.<sup>1</sup> The free-standing forms of the nouns in the modern (South) Munda languages, on the other hand, are frequently longer. A confounding array of different means of deriving the free forms of nouns from the noun roots are found in the individual Munda languages, often making reconstruction of specific proto-Munda forms impossible. One possible solution is the following: A minimally bimoraic shape appears to have been obligatory for free-forms of nouns in Proto-Munda and many of its daughter languages, which necessitated (and in Gta? continues to necessitate) the combining of the root element with some derivational process (in particular, prefixation, infixation, reduplication, compounding, or suffixation). This proposed Proto-Munda “Bimoraic Constraint on Free Forms of Nouns” [henceforth BC] can be roughly schematized as in (1).



where a mora  $\mu$  is linked both to any vocalic nucleus, a consonant in word initial position before another consonant, and, in some languages, a post-vocalic [ʔ] in final position (or a post-vocalic laryngeal), as in (2)



## 1 Problems in the reconstruction of the Proto-Munda noun

The elements used to fulfill the Bimoraic Constraint [BC] were varied in Proto-Munda. Prefixes were most commonly found, infixes less frequently, reduplication not uncommonly, and root-compounding on occasion, while suffixes were used in a very few restricted instances, particularly in North Munda languages, where the BC was less strong. Functionally the elements may have in certain instances been markers of noun-class (e.g. *\*k(V)-* 'animate'), or have no apparent semantic content at all, i.e. as far as we can tell, they were motivated at various stages purely by the need to fulfill the bimoraic constraint on free forms of nouns.

The BC is more active (or highly ranked in Optimality Theory terms) in some Munda languages than in others. For example, there are no underived, monosyllabic free forms of nouns at all in Gta? in the South Munda periphery, while in the Kherwarian North Munda language Santali, many monosyllabic, seemingly monomoraic nominal free forms are attested. Other Kherwarian languages show a greater degree of bimoraic or disyllabic noun forms.

While the range of formatives found with a given noun stem may vary greatly among the modern Munda languages, there are certain trends of derivation that can be observed. For example, a given Munda language may assign a particular noun-formative element as the default in the derivation of free forms of nouns; for example, the default derivational element in Gta? appears to be prefixed syllabic nasal.

In other languages, particular formatives are favored within definable, restricted sets of nouns. For example, nouns referring to animals tend to take the *\*kV-* prefix in the South Munda language Gutob. This patterning is non-random, has parallels throughout the AA family (Smith 1975), and may reflect some kind of archaic animal or animate classifier prefix. Another example is the favoring of a glottal stop infix in the derivation of body part terms in Sora-Gorum. Further research is required in order to determine whether this pattern is significant. Also, there is the potentially non-random patterning in Kharia of the prefix *kon-* in words referring to small animals or words roughly falling into a category of 'diminutive' or 'affectionate'; note in this regard the possible connection of this element to the word meaning 'child' as well as to the common classifier in Vietnamese *con*.<sup>2</sup>

However, in other instances, certain noun roots appear to favor a particular derivational element or process to create the free-standing form, for example the reduplication found in the word for

'turmeric' (albeit CVC- reduplication in South Munda and CV-reduplication in North Munda). Some correspondence sets seem to have very infrequent or isolated derivational elements in Munda, e.g. the sV- in 'stomach' (and possibly 'sambar deer' as well) or the pV- in the word for 'little bird'. Note that in North Munda there is a general tendency to eliminate prefixed elements and make more extensive use of suffixes or compounding. See examples in (3) below.

### (3) Selected Nouns in Munda languages (Zide 1999)

Gutob	Remo	Gta?	Kharia	Juang	gloss
titi	titi	titi	ti?	iti	'hand'
susuŋ	tiksūŋ	nco	--	ijjɪŋ/ŋ	'foot'
--	gisinra?e	gcæŋ	jinray	jinlae	'porcupine'
oŋger	ŋger	ngir	kɔŋg <sup>h</sup> er	kɔŋger	'yng man'
mɔd/?	mɔd	m-mwa?	mɔd	ɔmɔɪ/d	'eye'
gikil, kilɔ	kilɔ	ŋku	kiɾo(g)	kiɾɔg	'tiger'
ɔ?ɔn	ɔ?ɔn	V~-hV~	kɔnɔn	kɔnɔn	'child'
gusɔ?	gusɔd	gsu?	solo?	selag	'dog'
ɖa?	ɖa?	ndia?	ɖa?	ɖag	'water'
sulɔb	sulɔb	sla?	(uslo? 'earth')?	--	'tree'
sulɔj	sulɔi	slwe?	laej	--	'stomach'
sasaŋ	saŋsaŋ	ssia	saŋsaŋ	sa(ɾa)ŋ saŋ	'turmeric'
gidɛb	gidɛd	gri?	?	--	'frog'
	gisin	gsæŋ	sin̄koe/i, =sin [CF]	sen̄koe	'fowl'
suram	siram	sra	--	seɾam	'sambar deer'
gikin	ŋkui~	ŋkwi~/ŋ	--	kuin̄kar	'f-i-l, w.e.b."
gubon	gibe	gbe	bane/, -ai	banæ	'bear'
soʔl	suʔu	ncu	jol	ɔjɔn	'oil'
urɔi/urei	urai	n(d)rwe	kondroy	--	'fly'
piɾig	piɾi?	plæg	konthed	kontɛd	small bird
<u>Sora</u>	<u>Gorum</u>	<u>Kherw</u>	<u>Korku</u>		<u>gloss</u>
sʔi	siʔi	ti ~ tii	ti		'hand'
jʔeŋ	jʔiŋ	janga	naŋgà		'foot'
kəŋjin	ubaʔjin	jʔi-k jiki (H)	jikɾa		'porcupine'
oŋger-sij	iŋ-ger	--	--		'yng man'

mʔɔd, amad	mad	met/d, (-e-, -e-)	med		'eye'
kina	kul(a)	kul, kula (M)	kula		'tiger'
oʔon	aʔon	hɔn (o), hɔpɔn	kɔn		'child'
kənsod	kusod	seta	sita		'dog'
d(a)ʔa	daʔa	dak/gʔ/Ø	ɖaʔ/g		'water'
ənʔeb	--	--	--		'tree'
--	--	lac, lai(?)	laj		'stomach'
saŋsaŋ	saŋsaŋ	sasaŋ	sasaŋ (c-)		'turmeric'
kindud	--	--	ded-dàʔ		'frog'
kənsim <sup>3</sup>	(aŋ=oi)	sim	sim		'fowl'
kunsar	ki-sar	saram (H)	--		'sambar deer'
kupar	kipar	hɔpɔn <sup>4</sup>	kupkar		'f-i-l, w.e.b.'
kəmbud	kibud	bana	bana		'bear'
miŋol/=ŋol	--	--	--		'oil'
əroy	aroy	rɔ, roko	ruku		'fly'
ontid	porid	titid/r=	tit'id		small bird

#### (4) Select Munda Correspondences

Gutob	Remo	Gtaʔ	Kharia	Juang	Sora	Gorum	Kherw	Korku
Rdpl	Rdpl	Rdpl	-ʔ	*N-	-ʔ-	-ʔ-	Ø <sup>3</sup>	Ø
Rdpl	X-Y	*N-	--	*N-	-ʔ-	-ʔ-	-a	-a
--	*kV-X-Y	k-X	X-Y	X-Y	kən-X	Y <sub>2</sub> -X	Ø/-i	X-Y
oŋ- (< *N-)	*N-	*N-	kən-	kən-	*N/*kən ??	*N-	--	--
Ø	Ø	*N-	Ø	*N-	-ʔ-, a - /*N-	Ø	Ø	Ø
*kV-, -ɔ	-ɔ	*N-	-ɔg	-ɔg	-a	-a/Ø	*-a	-a
-ʔ-/*N-	-ʔ-/*N-	*N-	-n-	-n-	-ʔ-	Y-X, *N	Ø, -p-	Ø
*kV-	*kV-	*kV-	-l-	-l-	kən-	ku-/*kən	-a	-a
Ø	Ø	*N-	Ø	Ø	-a	-a	Ø	Ø
sV=-Ø-	sV=-Ø-	sV=-Ø-	??*sV-	--	*sV=-ʔ-	--	--	--
sV-	sV-	sV-	Ø	--	--	--	Ø	Ø
Rdpl CVC-	Rdpl CVC-	Rdpl CVC-	Rdpl CVC-	Rdpl CVC-	Rdpl CVC-	Rdpl CVC-	Rdpl CV-	Rdpl CV-
*kV-	*kV-	*kV-	--	--	kin-	--	--	X-Y
*kV-	*kV-	*kV-	X-Y	X-Y	kən-		Ø	Ø
X-Y	X-Y	X-Y	--	X-Y	*kən-	*kən-	X-Y	--
*kV-	*N-	*N-	--	X-Y	X-Y	X-Y	X-Y	X-Y
*kV-	*kV-	*kV-	X-Y	X-Y	*kən-	*kən-	-a	-a

-ʔ-	-ʔ-	*N-	Ø	*N-	Y-X	--	--	--
*V-/N-	*V-/N-	*N-	kɔn-	Ø	ɔ-/ *N-	a-/ *N-	Ø, -X- PL	X-PL
*pV-	*pV-	*pV-	kɔn-	kɔn-	*(k)ɔn-	*pV-	Rdpl (- Y)	Rdpl

[Sources: Ramamurti 1931; N. Zide field notes; A. Zide n.d.; Malhotra 1982; Biligiri 1965; Kullu 1981; Bhattacharya 1968; Pinnow 1960-ms.; Osada 1992; Deeney 1975; Campbell/Macphail 1954]

Key: X = CVC-root, Y = element used to form the free-standing compound, Y<sub>2</sub> = second element used in this construction, PL = plural

Note that certain subsets of semantically related nominal free forms seem to pattern together in individual Munda languages, for example, the words meaning ‘hand’ and ‘foot’ have the same pattern in, e.g., Gutob (with reduplication in both), Juang (with an etymological syllabic nasal prefix), and [Proto-]Sora-Gorum (with infixing glottal stop), with a different derivational element in each language/sub-group.

## 1 Evidence for the BC in other Austroasiatic languages

As the Munda languages do not form an isolated family, but rather constitute the westernmost branch of the large Austroasiatic [AA] family, in order to shed light on some of issues relating to the apparent BC active in Proto-Munda, we naturally turned to the broader comparative picture. As it turns out, problems similar to those facing the researcher of comparative Munda arise in the reconstruction of many other sub-families within Austroasiatic. This suggests that the Bimoraic Constraint may have been operative to derive free forms of nouns all the way back to Proto-Austroasiatic, which, as in Munda, has either been lost or preserved in an irregular manner in numerous daughter languages, both attested and reconstructed.

### *1.0 Problems in Katuic noun-form correspondences*

Within the Katuic branch of AA (5) , there are nearly 20 different patterns of correspondences seen listed in (6). These clearly

do not reconstruct back to twenty different initial sequences, but rather, such variation suggests that the BC was operative in deriving free forms of nouns in the history of Katuic, as in Munda, with different languages selecting different elements in the derivation of particular lexical items from an underlying root.

(5) Katuic Noun correspondences (Peiros 1996)

Bru	Kui	Pakoh	Katu	gloss
nci:ʔ		kəci:k		'comb'
phərce:l		ʔəcial	jəju:/il	'heart'
ce:m	ce:m		ʔəcim	'bird'
kəcah	kəcah-cah	kucah	kəcah	'charcoal'
ncɿj	nce:	nce: <sup>T</sup>	ncaj	'body lice'
ʔəca:	ca:-ʔaca:	ʔəcɔ:	ʔəcɔ	'dog'
kəhial	khi:l	kijjal		'bee, hornet'
kəhi:p	khəhe:p-hē:p	kəhe:p <sup>T</sup>	kəhip	'centipede'
ʔəha:m	ŋha:m	ʔəha:m	ʔəha:m	'blood'
kla:ŋ	kla:ŋ-kla:ŋ	kla:ŋ	kla:ŋ	'hawk, kite, eagle'
nluən <sup>B</sup>		kluən	cəlun	'calf, leg'
	məʔ-ʔa:məʔ		ʔəme:ʔ	'mother'
ka:n	ka:n	ʔkərnian <sup>T</sup>	ʔəka:n-kien	'child'
ruaj <sup>B</sup>	ruaj <sup>B</sup> -ʔa:ruaj <sup>B</sup>	riɾɔ:j	rəɾɔ:j (AD)	'fly'
ʔətɿj	te:-ʔa:te:	ʔəti:	tɿj	'hand'
	nta:ʔ	nta:ʔ	nta:k	'tongue'
	kəta:m-ta:m	ʔəta:m	ʔəta:m	'crab' (VN dam)
ntre:l	nthre:l	t(i)rial <sup>T</sup>	krial	'egg'
ŋha:ŋ	ŋha:ŋ	ŋha:ŋ	ŋha:ŋ	'bone'
ʔəjə:ŋ	dzi:ŋ	ʔji:ŋ	juŋ	'foot, leg'
ŋkim		nkim		'thumb'
	təkəm-kəm		təkə:m	'finger, toe'
ʔəka:	ka:-ʔa:ka:	ʔəka:		'fish'
ko:ŋ		ko:ŋ	ʔəkɔŋ	'father'
	wuaʔ-ʔa:wuaʔ		wɔ:k (AD)	'monkey'
			ʔbɔ:k	
təʔu:r	ŋʔo:r	ʔiŋo:r		'hornet, bee'
	ŋŋo:r <sup>B</sup> -ŋo:r <sup>B</sup>	< *ŋ(ə)ʔo:r		

unmarked = 'normal/clear/lax' voice

<sup>B</sup> =breathy

<sup>T</sup> = tense

Note also Bru səruaj 'horsefly'; Note Bru kɔ:ŋ ('of animal')

## (6) Select Katuic Correspondences

Bru	Kuy	Pakoh	Katu	Bru	Kuy	Pakoh	Katu
n-	--	k-	--	p <sup>h</sup> ər-	--	ʔə-	CV-
Ø	Ø	--	ʔə-	k-	k-	k-	k-
n-	n-	n-	n-	ʔə-	X=ʔə-X	ʔə-	ʔə-
ʔə-	ŋ-	ʔə-	ʔə-	n-	--	k-	cə-
--	X=ʔə-X	--	ʔə-	Ø	Ø	--	ʔə-Redpl
Ø	X=ʔa-X	CV-	CV-	ʔə-	X=ʔa-X	ʔə-	Ø
--	kə-X-X	ʔə-	ʔə-	Ø	Ø	Ø	Ø
n-	n-	Ø	k-	ŋ-	ŋ-	ŋ-	ŋ-
ʔə-	Ø	ʔ-	Ø	--	X=ʔa-X	--	Ø(/ʔ-)
tə-	ŋ-	*ŋ-	--				

## 1.0 Bahnaric Correspondences

Katuic is far from alone among eastern Austroasiatic languages in showing a range of nominal free forms analogous to the diversity seen in Munda. Many other branches exhibit a noteworthy degree of variability. For example, irregular correspondences can be found among various members of the Bahnaric family, see (7) and (8).

(7)

Bahnar	Jölöng	gloss
anah	tönah	'wood, tree'
köyaa	röyaa	'ginger'
röngaa	röngaa	'sesame'
tömoo	tömoo	'stone'

(Léger 1974: 124-5)

(8)

Bahnar	Cua	Chrau	Stieng	gloss
patuol		ntül	ttul	'anthill'
ʔdök	talôk			'monkey'
muh	muh	müh	tromüh	'nose'

(Gregerson, Smith, Thomas 1976: 393-7)

## 2.3 Palaung-Wa noun correspondences

When comparing languages of the Plang subgroup of Palaung-Wa (Paulsen 1992 : 210-3), a similarly confounding range of variation is found among the correspondence sets. See (9) and (10). As with Munda and Katuic, such a wide degree of variability is suggestive of the operation of the BC, allowing for the varied and individual selection of derivational formatives to derive the free forms of the nouns. As in Munda, a given language appears to prefer particular formative elements in order to fulfill the BC, e.g.  $ka\gamma^4$  in Shinman.

(9)

Kontoi	Shinman	Samtao	gloss
amhac <sup>1</sup>	ka $\gamma^4$ muik <sup>1</sup>	myc <sup>1</sup>	'ant'
ak <sup>h</sup> rak <sup>1</sup>	qhak <sup>1</sup>	krak <sup>1</sup>	'buffalo'
amoy <sup>2</sup>	ka $\gamma^4$ moi <sup>2</sup>	moi <sup>2</sup>	'cow'
kətam <sup>1</sup>	ka $\gamma^4$ tam <sup>1</sup>	tam <sup>1</sup>	'crab'
kənel <sup>1</sup>	ɛh <sup>1</sup>	kənia <sup>2</sup>	'chicken'
atep <sup>1</sup>	tiap <sup>1</sup>	tip <sup>1</sup>	'flea'
fa $\gamma^1$	---	kənfə $\gamma^1$	'gibbon'
ape $\gamma^2$	pe $\gamma^4$	pe $\gamma^1$	'goat'
a $\gamma$ uŋ <sup>1</sup>	ɔŋ <sup>1</sup>	ɔŋ <sup>1</sup>	'hornet'
apleŋ <sup>1</sup>	kliŋ <sup>1</sup>	piŋ <sup>1</sup>	'land leech'
kənlɪk <sup>1</sup>	lik <sup>2</sup>	kənlɛc <sup>2</sup>	'pig'
kənkəŋ <sup>2</sup>	kəŋ <sup>3</sup>	kənkəŋ <sup>2</sup>	'rat'
kənvay <sup>2</sup>	ka $\gamma^4$ vai <sup>3</sup>	avai <sup>2</sup>	'tiger'
na $\gamma$ wuk <sup>1</sup>		na $\gamma$ yk <sup>1</sup>	'chest'
rə $\gamma$ uh <sup>1</sup>	la $\gamma^4$ uL <sup>1</sup>	aluah <sup>1</sup>	'fat'
kətiŋtol <sup>1</sup>	ka $\gamma^4$ tiŋ <sup>2</sup>	tv $\gamma$ atəŋ	'navel'
kəvaŋ <sup>2</sup>	ka $\gamma^4$ vaŋ <sup>1</sup>	avaŋ <sup>2</sup>	'thigh'
ntak <sup>1</sup>	ka $\gamma^4$ tak <sup>1</sup>	ntak <sup>1</sup>	'tongue'
avoy <sup>1</sup>	oi <sup>1</sup>	o <sup>1</sup>	'fem. in-law'
akaŋ <sup>1</sup>	kuiŋ <sup>1</sup>	kɣŋ <sup>1</sup>	'father'
ata $\gamma^1$	ta $\gamma^1$	ata $\gamma^1$	'g.f.'
kəmi $\gamma^2$	ka $\gamma^4$ me $\gamma^2$	ame $\gamma^1$	'husband'
m̄mi $\gamma^2$	ka $\gamma^4$ me $\gamma^2$	konme $\gamma^1$	'man'
ama $\gamma^2$	ma $\gamma^2$	ma $\gamma^2$	'mother'
kənk <sup>h</sup> rəh <sup>1</sup>	--	pəkrih <sup>1</sup>	'single female'
maŋ <sup>2</sup>	ka $\gamma^4$ muiŋ <sup>3</sup>	amɣŋ <sup>2</sup>	'wife'
m̄puw <sup>1</sup>	ka $\gamma^4$ pɣŋ <sup>1</sup>	kənpun	'woman'
m̄pak <sup>1</sup>	ka $\gamma^4$ puwk <sup>2</sup>	apɣk <sup>1</sup>	'bridge'

ə = subscript +



(10)

<u>Kontoi</u>	<u>Shinman</u>	<u>Samtao</u>	<u>Kontoi</u>	<u>Shinman</u>	<u>Samtao</u>
a-	kaʔ=	Ø	a-	Ø	Ø/--
kə-	kaʔ=	Ø	kən-	Ø	kən-
Ø	--	kən-	a(?)-	Ø	Ø
a-	k-	Ø	-ʔ-	--	-ʔ-
kən-	kaʔ=	a-	kə-...-tol	kaʔ=	tʏ-ʔ-a-
-ʔ-	laʔ=	a-	a-	Ø	a-
ŋ-	kaʔ=	ŋ-	m-	kaʔ=	kən-
kə-	kaʔ=	a-	m-	kaʔ=	a-
kən-	--	pə-			

## 2.4 Nicobarese Correspondences

To be sure, most, if not all, subgroups of Austroasiatic show disturbing inconsistencies in correspondences among free forms of nouns, despite the presence of a cognate root element across the languages of the subgroup. Compare in this regard, the forms of 'hand' found in the various languages of the Nicobarese subgroup in (11).

(11)

<u>Central</u>	<u>Car</u>	<u>Shom Pen</u>	<u>Teressa</u>	<u>gloss</u>
kane-tai	el-ti:	noai-ti:	mòh-ti:	'hand'

(Man 1975 [1888-9])

## 1.0 Aslian Correspondences

Members of the Aslian branch of Austroasiatic also show a certain degree of variation with regards to the elements used in deriving free forms of nouns (Benjamin 1976: 102ff.). Many Aslian languages seem to have lost the BC altogether, but others seem to have preserved it. A few examples from Aslian languages are offered in (12).

(12)

<u>Sn, Sa<sup>6</sup></u>	<u>LJ, LY</u>	<u>TM</u>	<u>SM, I/II</u>	<u>MM, SB, Sl, Tq</u>	<u>gloss</u>
ko:n	kəwən	kəwət	kəno:n	kənən	'child'

<u>Ks, KB, Je, BN, CW</u>	<u>Mr</u>	<u>JH</u>	<u>gloss</u>
wəŋ	kəwəl	ʔewaʔ	'child'

<u>Sn</u>	<u>Sa</u>	<u>LY</u>	<u>Tm</u>	<u>Sm.i</u>	<u>Sm.ii</u>	<u>JH</u>	<u>MM</u>	<u>SB,S1</u>	<u>Tq</u>	<u>gloss</u>
ʔacɛh	co:ʔ	ʔacɛʔ	cəwɔʔ	cɔ:ʔ	coʔ	cuwo:ʔ	caw	cɔ:h	cɔw	'dog'

<u>Sn</u>	<u>Sa</u>	<u>JH</u>	<u>gloss</u>
ʔəŋkɔŋ	ʔəŋkɔŋ	kərakɔŋ	'male'

<u>CW</u>	<u>Sn, Tm, Sm.i, Sm.ii</u>	<u>Sa, LY, JH</u>	<u>MM</u>	<u>SB, S1</u>	<u>Tq</u>	<u>gloss</u>
beʔ	ba:ʔ	baʔ	beʔ	baba:h	babah	'rice'

## 2.6 Pearic Correspondences

Pearic languages (Headley 1978: 86) also show a limited degree of variation in free forms of nouns. Compare in this regard the words for 'shrimp' seen in (13) in various members of the Pearic sub-group.

(13)

<u>Somray</u>	<u>Somree.</u>	<u>Chong həəp</u>	<u>Chong lɔɔ</u>	<u>gloss</u>
mpí:h	kəpih	kəpi:h	kəpi:t	'shrimp'

As in Munda, Katuic, and Palaung-Wa, individual Pearic languages appear to have generalized a particular formant to fulfill the BC on nominal free forms. Compare in this regard the Chong-Song correspondences in (14), where Chong seems to have generalized the *\*k(V)-* prefix, corresponding to at least three different prefix elements in Song (Diffloth 1989: 149).

(14)

<u>Chong</u>	<u>Song</u>	<u>gloss</u>
kəlaʔ	khla'a	'leaf'
kəlaʔ	səla:	'thorn'
kəpha:	ləpha:	'tortoise'

Note also Chong *kakhoo* vs. Samre *rəkhàw* 'husked rice' (Pou and Martin 1981: 43).

(cf. Talan [Katuic] *tapa:*, Vietnamese *ba-ba*)

## 2.7 Monic Correspondences

Monic languages are similarly not exempt from the effects of the purported BC under investigation (Huffman 1990: 58-83). For

example, most of the Monic forms meaning ‘eel’ or a particular species of ‘eel’ suggest a \**kVn-* prefix, but the Central Nyah Kur dialects show a reduplicated form (Diffloth 1984: 69-97). Another good example of the variable means of deriving free forms of nouns in Monic languages is seen in the word for ‘kite’; see (16).

(15)

Burmese Mon	Thai Mon	Nyah Kur	gloss
kon	kon-ŋac	kuan-kuan	‘child’

(16)

<i>Mon(Ro)</i>	<i>Mon (Rao)</i>	<i>Mon(Thai)</i>	<i>gloss</i>
həcem	həcem	həcɛem	‘bird’
ʔəneiaŋ	k/ʔəneiaŋ	ʔ/kəniəŋ, həneiaŋ	‘kite’
ʔəyaoʔ	k/ʔəyaoʔ	k/ʔəyaoʔ	‘worm, maggot’
ʔəchao	k/ʔəchao	(ʔə)chao	‘red ant’
	hələŋ	hələyŋ	‘eel’
<i>NyahKur(ʔ)</i>	<i>NK(C)</i>	<i>NyK(S)</i>	<i>gloss</i>
ŋciam	ŋciam	kəŋciam	‘bird’
(ʔo)ŋlijaŋ	ləŋlijaŋ	liŋlijaŋ	‘kite’
ŋc(h)uuʔ	ŋchuuʔ	kənsuuʔ ~ kəŋchuuʔ	‘worm, maggot’
chaw	chaw	saw ~chaw	‘red ant’
ntoŋ	thunthoon	kənthoon	‘eel’

## 2.8 Khmeric Correspondences

The Khmeric branch of Austroasiatic also shows irregular correspondences, for example between Old Khmer and modern Khmer (Pou & Martin 1981: 16, 18, 28). See a brief set of examples in (17).

(17)

Old Khmer	Modern Khmer	gloss
cincaañ	añcaañ	' <i>gmelina asiatica</i> '
daṅdraañ ~ di-	kand(r)aamṇ <sup>(a)</sup> [hae]	' <i>polygonum barbatum</i> '
kañchet	kañchaet	' <i>neptunia oleracea</i> '

There are also numerous irregular correspondences between Khmer and other Austroasiatic languages with respect to the element used to fulfill the BC, e.g. Old Khmer *phcik* vs. Bahnar *kəcik* '*shorea obtusa*' (Pou & Martin 1981: 40).

## 2.9 Viet-Muong Correspondences

Even Viet-Muong languages show variation that is suggestive of a formerly active BC in their history. For example, there are a number of words that appear in a reduplicated form in Muong that correspond to simplex forms or compound forms of some type in Vietnamese (18).

(18) (Nguyen Van Khon 1987; Sokolovskaja/Nguyen 1987)

Muong	Vietnamese	gloss
dan <sup>3</sup> dan <sup>3</sup>	dán đi	'bat'
dan <sup>3</sup> tlan <sup>3</sup>	trán	'forehead'
ka <sup>1</sup> tlol <sup>2</sup>	một giống gà rùn	'grouse, blackcock'

Irregular correspondences of Thavung and Ruc to other Viet-Muong languages are also suggestive of the operation of the BC in the history of this branch of Austroasiatic.

(19) (Hayes 1992: 222)

Vietnamese	Muong	Ruc	Thavung	gloss
tóc <sup>1</sup>	thắc <sup>1</sup>	ũsúk <sup>1</sup>	sok <sup>1</sup>	'hair'
--	trốc <sup>1</sup>	kũluòk <sup>1</sup>	?ɔɔk <sup>1</sup>	'head'
răng <sup>1</sup>	thăn <sup>1</sup>	kàsăn <sup>1</sup>	ksan <sup>1</sup>	'tooth'
ruồi <sup>1</sup>	ruôi <sup>1</sup>	mũrɔy <sup>1</sup>	mɔyh <sup>2</sup>	'fly'

(20) (Ferlus 1974: 73)

ThavUng	Kha bô	A-rem	Vietnamese	gloss
atak	atak	atek	đất	'land'
akol	klo	t'lo		'star'

## 2.10 Khasic Correspondences

Among the languages of the Khasic branch of Austroasiatic, there are also instances of irregular, varied correspondences. The BC seems to be very strong in Khasi, where numerous free forms of nouns include some additional element. Quite common in Khasi proper as well as in Synteng, is the velar prefix. This is found on a wide range of animal names, kin terms, and body parts. However, in other languages of the branch, these elements occur less frequently, either replaced by a different element, or lacking altogether. Compare the forms listed in (21) in this regard (Fournier 1974: 86-92).

(21)

<u>Khasi</u>	<u>Lyngngam</u>	<u>Synteng</u>	<u>Amwi</u>	<u>gloss</u>
ksew	ksu:/'su:	ksaw ~kswa	ksiá	'dog'
sim	sim	sim		'bird'
khmat	kh'mat	khmat	ma:t	'eye'
khmut	leo-'mut	khmut	mur-kon	'nose'
<u>Lakadong</u>	<u>Mynnar</u>	<u>War</u>		<u>gloss</u>
ksaw	ksow	ksià		'dog'
	ksem	ksem		'bird'
ma:t		ma:t		'eye'
mur-kon		myrkon		'nose'

## 3 Evidence from Alternations: Derived Bimoraic Free Forms and Monosyllabic Combining Forms

Further evidence that is suggestive of the operation of the BC in the history of Munda and Austroasiatic comes from variation seen in nominal and verbal compounds generally consisting of two elements. One of these is a shorter, combining form of a noun which has a corresponding longer free form that fulfills the BC. Alternations of this type are common in Khasi, various South Munda languages (and Proto-South Munda), and Nicobarese.

3.1 Khasi

In Khasi, the free form of the word for ‘hand’ is *kṭi*, but this appears as simply *ti* when used as a combining form; for examples of this and other alternations, see (22-3).

(22) Khasi alternations between derived free form and unprefixing combining form

*kṭi* but *tiipdeŋ* ‘middle finger’ (Rabel 1961: 44)

*khmat* but *matli?* ‘white of eye’  
also *ʔiimat* ‘eye’ < see-eye/face (Rabel 1961: 149)

*khnaay* ‘mouse. rat’ but *naaysaaw* ‘small red hill mouse’

(23)

*kpa*, *kmi(e)* ‘father, mother’ (non-vocative) vs.

*ʔi paa ʔi mey* address term used by children to parents  
‘daddy/mommy’ (Rabel 1961: 49)

3.2 Munda

In various South Munda languages, nominal compounds and verbs with noun-incorporation use a set of short combining forms which generally correspond to a bimoraic free form of the noun (Mahapatra and Zide 1972). This is most highly developed in the SM language Sora as in (24), but is characteristic of all the South Munda languages to some degree.

(24) Sora Full Forms and Combining Forms (Starosta 1992: 85-86; Ramamurti 1931: 69ff.)

<u>Full Form</u>	<u>Combining Form</u>	<u>Gloss</u>
<i>əsu</i>	<i>=su</i>	‘illness’
<i>aŋgaj</i>	<i>=gaj</i>	‘moon’
<i>daʔa</i>	<i>=da</i>	‘water’
<i>jiʔi</i>	<i>=ji</i>	‘tooth’
<i>usal</i>	<i>=sal</i>	‘skin’
<i>kumbul</i>	<i>=bul</i>	‘rat’
<i>kinsod</i>	<i>=sod</i>	‘dog’

<i>saŋsaŋ</i>	= <i>saŋ</i>	'turmeric'
<i>jəno</i>	= <i>jə</i>	'broom'
<i>kənuŋ</i>	= <i>kuŋ</i>	'razor'
<i>siʔi</i>	= <i>si</i>	'hand'

*lam-si-t-am* [bow-hand-NONPAST-2] 'I bow to your hand(s)' (Ramamurti 1931: 43).

### 3.3 Nicobarese

Similar alternations between longer, derived free forms, and shorter, seemingly underived combining forms are found in Nancowry Nicobarese as well; see (25). As in South Munda, the short forms also appear in both nominal compounds and incorporated into verbs.

(25) Nancowry full and combining forms (Radhakrishnan 1981)

*ʔu(ál)mát* 'eye' vs. *kap-mat* 'imagine', *halepimatri* 'examine something.'

## 4 Conclusions

In Proto-Munda, and many of its daughter languages, a Bimoriac Constraint on free-forms of nouns appears to have been (or to continue to be) active, requiring the predominantly monosyllabic noun roots to undergo some elaboration process, either prefixation, infixation, compounding, suffixation, or reduplication. A similar problem arises in the reconstruction of numerous Austroasiatic subgroups.<sup>7</sup> This suggests the Proto-Munda feature may be an old one in the Austroasiatic language family as a whole. The Bimoriac Constraint on nominal free-forms may be preserved or lost in an irregular manner in individual Austroasiatic languages, both attested and reconstructed. In the case of (partial) preservation, a given language may have generalized a particular derivational element in this to fulfill the constraint e.g. *kaʔ* in Shinman (Palaung-Wa), *\*k(V)-* in Chong (Pearic), or *\*N-* in Gta? (South Munda).

Furthermore, the Khasi-Munda-Nicobarese parallels addressed in 3 above strongly suggest that alternations between longer free forms of nouns fulfilling the Bimoriac Constraint and shorter combining forms of nouns used in compounding and incorporation also may be quite old in the family, perhaps even being characteristic of the Proto-Austroasiatic ancestor language itself. Resolving this issue, however, must await further research in comparative Munda and

comparative Austroasiatic verbal and nominal morphology and phonology.

Notes

<sup>1</sup> There are a few roots in Munda that appear to be disyllabic in form. Some of these have laryngeal elements originally, e.g. the word for 'thigh'. Note that in Sora, some disyllabic forms, including certain loans, have no corresponding monosyllabic combining form and consequently cannot be incorporated or compounded. As these forms always fulfill the BC, but show no significant alternations, they are not exceptions to our claim that Proto-Munda noun roots were minimally bimoraic, but rather seem to constitute a minor form class of the language.

<sup>2</sup> Additionally, there may have been a historical conflation/contamination of two originally distinct morphemic sequences in a given Munda language, e.g. a prefix *\*kon-* <coming originally from 'child'> and *\*kV-n-*, an *\*n*-infix form attached to the previously mentioned *\*kV*-prefixed form or a *\*kV*- prefix attached to form already bearing a syllabic nasal prefix *\*N-*.

<sup>3</sup> Note also Sora *kənreŋim*, morphemically *kən-reŋ-im* the last syllable perhaps coming from *\*sim* with *s*-loss.

<sup>4</sup> Note also Ho *honyar*, Mundari *honyar* ~ *hoñear*.

<sup>5</sup> 'hand' has a long vowel in Asuri and Ho. These forms therefore satisfy the BC.

<sup>6</sup> List of Aslian language Abbreviations

KS	Kensiu	KB	Kentagbong	Je	Jehai
Mt	Mintil	BN	Bateq Nong	CW	Che' Wong
BD	Bateg Deq	Sa	Sabum	LJ	Lanoh Jengjeng
Tm	Temiar	Sm.i/ii	Semai I, II	JH	Jah Hut
SB	Semaq Beri	Sl	Semelai	Tq	Temoq
Mr	Mendriq	Sn	Semnam	LY	Lanoh Yir
MM	Mah Meri				

<sup>7</sup> As was pointed out in the discussion following this paper (Madison, WI May, 2000), a similar constraint appears to be characteristic of various stages or subgroups of Hmong-Mien languages as well (M. Ratliff, p.c.). The BC is also apparently at the heart of the Minimal Word theory (M. Macken, p.c.).



---

## References

- Banker, E. M. 1964a. Bahnar Affixation. *Mon-Khmer Studies I*: 99-118.
- 1964b. Bahnar Reduplication. *Mon-Khmer Studies I*: 119-134.
- Bars, Rev. E. 1973. *Khasi-English Dictionary*. Shillong: Don Bosco.
- Benjamin, G. 1976. Austroasiatic Subgroupings and Prehistory in the Malay Peninsula. In Jenner et al. [eds.], p. 37-128.
- Bhattacharya, S. 1968. *A Bonda Dictionary*. Poona: Deccan College.
- Biligiri, H. S. 1965. *Kharia: phonology, grammar, vocabulary*. Poona: Deccan College.
- Cohen, P. 1965. Presyllables and Reduplication in Jeh. *Mon-Khmer Studies II*: 91-103.
- Costello, N. A. 1965. Affixes in Katu. *Mon-Khmer Studies II*: 29-56.
- 1991. *Nôôq Paraaq Katu [Katu Dictionary]*. Manila: SIL Thailand.
- Deeney, J. 1975. *Ho Grammar and Vocabulary*. Chaibasa: Xavier Ho Publications.
- Diffloth, G. 1976. Jah-Hut, an Austroasiatic language of Malaysia. In *South-East Asian Linguistic Studies*. *Pacific Linguistics C-42*: 73-118. Canberra.
- 1980. *The Wa Languages*. *Linguistics in the Tibeto-Burman Area 5.2*. Berkeley.
- 1984. *Dvaravati Old Mon and Nyah Kur*. Chulalongkorn: University of Press.
- 1989. Proto-Austroasiatic Creaky Voice. *Mon-Khmer Studies XV*: 139-54.
- Ferlus, M. 1974. Les langues du groupe Austroasiatique-Nord. In *ASEMI 5.1*: 39-68. Paris.
- Fournier, A. 1974. Les Khasi, une population mon-khmer de l'Inde. In *ASEMI 5.1*: 79-96.
- Gradin, D. 1965. Consonantal Tone in Jeh Phonemics. *Mon-Khmer Studies II*: 73-90.
- Gregerson, K, K. Smith and D. Thomas. The Place of Bahnar Within Bahnaric. In Jenner et al. eds., part I, pp. 371-406.
- Hayes, La Vaughn H. 1992. On the track of Austric: part I. In *Mon-Khmer Studies XXI*: 143-78.
- Headley, R. K. Jr. 1978. *An English-Pearic Vocabulary*. *Mon-Khmer Studies VII*: 61-94.
- Henderson, E. J. A. 1976. Vestiges of Morphology in Modern Standard Khasi. In Jenner, P. N., L. C. Thompson and S. Starosta. [eds.]. part I, pp. 477-522.

- 
- Huffman, F. E. 1990. Burmese Mon, Thai Mon, and Nyah Kur: A Synchronic Comparison. *Mon-Khmer Studies* XVI-XVII: 31-84.
- Jacob, J. 1993. *Cambodian Linguistics, Literature and History. Collected Articles*. D. A. Smythe [ed.]. London: SOAS.
- Jenner, P. N. 1982. *A Lexicon of Khmer Morphology*. *Mon-Khmer Studies* IX-X. Honolulu.
- Jenner, P. N. L. C. Thompson and S. Starosta. [eds.] 1976. *Austroasiatic Studies, Part I & II*. *Oceanic Linguistics Special Publication* no. 13. Honolulu: University of Hawaii Press.
- Kullu, P. 1981. *Khaṛiyaa vyaakaraṇ evaṃ sāṅksipt śabdakoś*. [Kharia grammar and short lexicon]. Dharmik Saahtya Samiti. Ranchi.
- Léger, D. 1974. Vocabulaire comparé et recherche du vocabulaire dentaire Bahnar-Jolong. In *ASEMI* 5.1: 123-32.
- Macphail, R. M. 1954. *Campbell's Santal-English Dictionary*. 3<sup>rd</sup> ed. Benegaria: Santal Mission Press.
- Mahapatra, K. and N. H. Zide. Nominal Combining Forms in Gta?. *Indian Linguistics* 33: 79-102.
- Malhotra, V. 1982. The structure of Kharia: a study in linguistic typology and change. J Nehru Univ. PhD dissertation.
- Man, E. H. 1975. [1888-9]. *A Dictionary of the Central Nicobarese Language*. Delhi: Sanskaran Prakashak.
- Means, N & P. B. Means. 1986. *Sengoi-English English-Sengoi Dictionary*. Toronto.
- Miller, J. D. 1964. Word Classes in Brôu. *Mon-Khmer Studies* I: 41-62.
- Nacaskul, K. 1978. The Syllabic and Morphological Structure of Cambodian Words. *Mon-Khmer Studies* VII: 183-200.
- Nguyen, V. Kh. 1987. *Tự-Diên Anh-Việt*. English-Vietnamese Dictionary. Glendale, CA: Dainam Publishing.
- Osada, T. 1992. *A reference grammar of Mundari*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Paulsen, D. A 1992. Phonological Reconstruction of Proto-Plang. *Mon-Khmer Studies* XVIII-XIX: 160-222.
- Peiros, I. 1996. *Katuic Comparative Dictionary*. *Pacific Linguistics* C-132. Canberra: ANU.
- Pinnow, H-J. 1960-ms. Beiträge zur Kenntnis der Juang-Sprache. Unpublished-ms. Berlin.
- Pogibenko, T. G. and Buy Kh. Th. 1990. *Język Ksingmul*. Materialy

- 
- sovetsko-v'etnamskoj lingvističeskoj èkspedicii 1979 goda. Moscow: Nauka.
- Pou, S. and M.-A. Martin. 1981. Les noms de plantes dans l'épigraphie vieux-khmère. In ASEMI XII/1-2: 3-74.
- Rabel, L. 1961. Khasi: A Language of Assam. Baton Rouge: Louisiana State University Press.
- Radhakrishnan, R. 1981. The Nancowry Word. Edmonton: Linguistic Research, Inc.
- Ramamurti, G. V. 1931. A Manual of the So:ra: (or Savara) Language. Madras: Govt. Press.
- Smalley, W. A. 1961. Outline of Khmu? Structure. New Haven: AOS.
- Smith, K. D. 1972. A Phonological Reconstruction of Proto-North-Bahnaric. Ukarumpa, PNG: SIL.
- . 1975. The velar-animal prefix relic in Vietnam languages. *Linguistics in the Tibeto-Burman Area* 2.1: 1-18.
- . 1992. The -VC Rhyme Link between Bahnaric and Katuic. *Mon-Khmer Studies* XVIII-XIX: 106-59.
- Sokolovskaja, N. K. and Nguyen V. T. 1987. Jazyk Muong. Materialy sovetsko-v'etnamskoj lingvističeskoj èkspedicii 1979 goda. Moscow: Nauka.
- Starosta, S. 1992. Sora Combining Forms and Pseudo-Compounding. *Mon-Khmer Studies* XVIII-XIX: 78-105.
- Svantesson, J.-O. 1983. Kammu Phonology and Morphology. Malmö (Lund): CWK Gleerup.
- Thomas, D. 1990. The instrument locative and goal affix -N- in Surin Khmer. In *Mon-Khmer Studies* XVI-XVII: 99-108.
- Wallace, J. M. 1965. Katu Personal Pronouns. *Mon-Khmer Studies* II: 19-28.
- Watson, S. K. 1964. Personal Pronouns in Pacoh. *Mon-Khmer Studies* I: 81-98.
- Zide, A. n.d. A Gorum-English Lexicon. Unpublished-ms. Chicago.
- Zide, N. 1999. Noun Formatives in South Munda: Implications for South Munda Subgrouping. Presented at SALA 1999, Urbana, IL.
- . Field Notes on Korku.
- . Field Notes on Gutob, Gta?.

