The origins of nominal affixes in MSEA languages: convergence, contact and some African parallels

Paper presented at

Mainland Southeast Asian Languages: The State of the Art in 2012

29 November - 1 December 2012

Max Planck Institute for Evolutionary Anthropology, Leipzig

and subsequently revised

Roger Blench
Kay Williamson Educational Foundation
8, Guest Road
Cambridge CB1 2AL
United Kingdom
Voice/ Ans (00-44)-1223-560687
Mobile worldwide (00-44)-(0)7847-495590
E-mail rogerblench@yahoo.co.uk
http://www.rogerblench.info/RBOP.htm

Cambridge, Thursday, 23 May 2013

TABLE OF CONTENTS

1. INTRODUCTION **1**
2. WORD STRUCTURE IN AUSTROASIATIC **3**
3. WORD STRUCTURE IN SINO-TIBETAN **8**
4. CONTACT, BORROWING AND METATYPY **11**
5. PARALLEL PROCESSES IN AFRICAN LANGUAGES **13**
6. A HISTORICAL SCENARIO **16**
7. CONCLUSIONS **17**
REFERENCES **17**

TABLES

Table 1. An Austroasiatic root for 'bear' with variable prefixes 3
Table 2. Animal names in Nafra Miji 4
Table 3. Hruso animal names with S- prefix 4
Table 4. Miji and Bangru body parts with an m- prefix 5
Table 5. ‘Buffalo’ *#k.r.pu.y* in Austroasiatic 6
Table 6. The *#kVla* root for ‘tiger’ in SE Asian languages 7
Table 7. The root C.ni.C(C) for 'two' in Tibeto-Burman 8
Table 8. Affix movement and loss in the evolution of Chinese 10
Table 9. ‘Bear’ in Tibeto-Burman languages 10
Table 10. Ut-Ma’in noun class prefixes 14
Table 11. Proto-Heiban noun-class prefixes 15
Table 12. Erosion and restructuring of CV nominal prefixes in Benue-Congo 16

FIGURES

Figure 1. Cyclical renewal of affixes in SE Asian languages 12

PHOTOS

Photo 1. Hruso shaman, Thrizino 4
Photo 2. Mangosteen 12
Photo 3. Heiban houses, Nuba Mountains 15

1. Introduction

The languages of the world characteristically have morphological strategies both to classify nouns and signify to both speakers and hearers aspects of their semantics. These strategies can be broadly divided into two widespread categories, noun-classes and numeral classifiers. A third system, characteristic of phyla such as Indo-European and Afroasiatic, is the sex-gender system, classifies nouns through notional male/female oppositions, but which is often not informative about semantics. Noun classes are often referred to as 'genders' but this is a confusing terminology as it conflates a genuine semantic categorisation with sex-gender systems, which broadly do what they say on the tin. Global overviews such as Aikhenvald (2000) sometimes conflate these two systems. These categories are not watertight; languages can display aspects of all these, and erosion of one morphosyntactic category can lead to the partial or complete evolution of another. For example, the Kana language in the Niger Delta of Niger is historically a Niger-Congo language descended from languages with rich noun-class systems. However, along with the other members of the Ogoni group, it has almost entirely lost this aspect of its morphology and developed instead a system of numeral classifiers (Ikoro 1996). Similarly, Krongo, one of the Kadu languages, a branch of Nilo-Saharan, appears to have nominal affixes, although these are partly fossilised and unproductive. They have no semantic assignments, but Krongo has instead adopted or developed a sex-gender system (Reh 1985, 1994; Blench 2006).

Languages are, broadly speaking, conservative and phyla can be characterised by particular strategies. So the great majority of Niger-Congo languages have noun-classes or nothing; numeral classifiers or sex-gender systems rarely develop. Afroasiatic languages exhibit sex-gender throughout the phylum and indeed a highly conservative morphology. Austroasiatic, Austronesian, Hmong-Mien and Sino-Tibetan languages all have numeral classifier systems which do not show agreement. It is usually considered that noun classes are absent in SE Asia, although the nominal prefixes of Hmong-Mien are characterised as 'weakly classifying' (Ratliff 2010: 267). There is a third system of marking nominal semantics, the rather ill-defined 'class-terms' which are found in Lao and Thai (Enfield 2007). These consist of obligatory accompaniments to nouns marking their semantics, such as Thai marking all fish with a preceding *pla*, corresponding to Lao *paa*. English has a fragmentary system of this type in that the names of birds and fish are sometimes accompanied by the term itself ('blackbird', 'mutton-bird', 'dogfish', 'catfish'). Class terms are always etymologically transparent. Exactly how widespread they are in SE Asia is unclear since their description is often conflated with numeral classifiers. Concordial noun classes are widespread, though far from universal, in the Papuan and Australian language areas (Harvey & Reid 1997), and appear occasionally in New World phyla, in keeping with their great diversity.

This paper argues that the morphology of MSEA languages suggests that noun classes were formerly significantly more productive. Typical noun class languages consist of a root and an affix which can be prefixed, suffixed, infix or appear as an or circumfix. In rare cases, languages exhibit double-affixing, two separate affixes which alternate according to distinct rules. Examples of such languages are Bassari on the Togo-Ghana borderland and the Tivoid languages of SE Nigeria (Greenberg 1977). As part of the erosion of such systems, various types of fusion can occur, but the original morphemes are usually reconstructible. The affix in principle has a semantic assignation, and in some languages this is clear, whereas elsewhere the semantics can be opaque. Affixes frequently alternate; thus singulars can have one or two marked plurals. In Nilo-Saharan it is often considered that the 'middle' is the unmarked term and a singulative and a plurative can be formed from the root (Dimmendaal 2000). But this is not a necessary requirement of a noun class language; in Niger-Congo the m- class for mass nouns is always an unpaired class (Greenberg 1963; Blench 1995). Another important aspect of classic noun-class languages is concord, agreement between the nominal affix and affixes on other parts of speech in the same clause or sentence. Most typically concord is with adjectives, but verbs, pronouns and other particles can all enter the picture. Agreement can be alliterative, i.e. the same or segmentally related material is repeated, or non-alliterative, i.e. the relationship is regular without segmental copying.

In SE Asian language phyla, word structure is often described as 'sesquisyllabic' (Matisoff 1973). The intended meaning is that words have major and minor syllables, i.e. an iambic structure. The major syllables are in the stem and the minor syllable a prefix, generally C or CV. Since the –V is often represented

orthographical with a mid-central vowel, it may be that it is not realised segmentally. Minor syllable prefixes are either lost or optional in many languages, and they seem to change in ways that do not suggest phonological shift but affix substitution. Austroasiatic, Sino-Tibetan and Hmong-Mien are of this type. In Austronesian languages, the tendency is for the prefix to have a (C)V format and to be conserved. In other words, once a prefix has been fused to a stem, it is retained, from Taiwan to New Zealand, as it were. Daic languages are typically CV(C) and except in rare cases no longer retain the minor syllable.

Minor syllables thus have the appearance of optional prefixes in many languages. Anderson (2004) observes that in Munda the final syllable is the 'stable, meaning-associated element' while the prefixed syllables are unstable and cannot be assigned a meaning. Intriguingly, Blust (1988) also identifies an apparently similar system in Austronesian, where roots seem to retain a cross-language basic meaning, but are preceded by a variety of CV prefixes which transform the meaning in individual languages. Extended examples can be seen in 'roots' section of the online Austronesian Comparative Dictionary[1]. Blust considers this is an example of phonosemantic association, similar to phonaesthemes identified elsewhere in the world (e.g. sl/gl in English). However, as Sagart (2011) observes, the system has striking similarities to the MSEA structures identified here.

It is certainly the case that the minor or prefix syllables have no obvious semantic assignations, and in no SE Asian language do they show concord. But to assume that they have 'no meaning' suggests a curious model of language. A general postulate of morphology is surely that the elements of words either do or formerly have had meanings. These can be obscured over time, but the task of linguistics is surely to tease them out. In the case of the prefixes of MSEA languages, the fact that they vary dynamically from one language to another is surely a reflection of their significance for speakers. They surely cannot be simply euphonious noise.

Yet in other ways they do have the appearance of an unproductive noun-class system, and in this they resemble Nilo-Saharan. Nilo-Saharan languages seem to show a wide range of affixes that suggest there was formerly a fully functioning nominal marking system, but synchronically, no language shows a system as productive as those in Niger-Congo. Affixes certainly change to mark number, but alliterative concord is unknown. This system is stable, and Nilo-Saharan shows no sign of eliminating this unproductive morphological baggage and heading towards a SE Asian type system.

Austroasiatic and many branches of Sino-Tibetan[2] have a common word structure where the root is preceded by a C- prefix. Although C- prefixes may have semantic correlates, this is not consistent. In different languages the prefix may disappear or be substituted, while the root remains static (cf. examples in Matisoff 2003). The C- prefix can sometimes be incorporated into the stem, and a new prefix added, leading to complex initial sequences. Additional evidence for this is drawn from the typical pattern of pronouncing initial consonant sequences as individual segments; thus 'spr' in Austroasiatic and Sino-Tibetan is pronounced s.p.r, and is not a cluster as in Indo-European. Unlike true noun-class languages elsewhere in the world, no number-marking is implied, nor is there any trace of concord with adjectives or pronouns. These similarities between the two phyla are rather perplexing, as few historical linguists consider them to be related. Globally, such systems are extremely rare, and for them to exist in direct geographical proximity is unlikely if they arose independently. However, clear examples of common lexemes of any time-depth are few[3], and these are often shared with other regional phyla such as Daic and Hmong-Mien. Constructing a historical scenario to account for this structural convergence is not obvious; the likely Urheimats of these two phyla are far apart.

[1] It is striking that the majority of Blust's examples focus on Western Malayopolynesian, especially the Philippines. It is as if the system is completely dropped in Oceanic.

[2] I am aware this is a controversial term and that some prefer Tibeto-Burman. Whatever the case, Sinitic languages are in consideration, as the evidence for the type of canonic form described here is well attested in Old Chinese.

[3] Benedict (1992:4) says, 'there is little evidence of any borrowing of lexical items of 'core' type by TB/ST from AA/MK'. However, he does go by giving some striking examples from kin terms, as also the animal names 'hawk' and 'tiger' (cf. Table 6).

This paper[4] will describe the features of word structure in Austroasiatic and Sino-Tibetan that appear to be convergent and suggest how they might have arisen. It will argue that such features are transitional in terms of the evolution of true noun-classes and introduces a typological parallel from West Africa. The hypothesis is that the affix system originates from frozen nominal classifiers (and noun class-terms) and that as the system is renewed, these prefixes co-exist together with productive classifiers. It will examine possible borrowing scenarios and suggest that these can be detected, but are inadequate to explain the diachronic morphology. The paper will suggest the origins of this pattern lie in a period when the region was far more phyletically diverse and where an original pattern of CV.CV(CV) characterised languages of the region.

2. Word structure in Austroasiatic

Word structure in Austroasiatic (at least for nouns) seems to consist of one or more optional C(V) prefixes, a CV(CV) stem and a C suffix, often weakened to a glottal stop or deleted. The optional C(V) prefixes are often referred to as a 'pre-syllable' in the literature. Many nouns may well have had a labial or palatal approximant in final position. This has strong tendency to be incorporated into the stem, and if it represents a different prosody from the vowel of the stem, then the synchronic output may be either a diphthong or a long vowel.

There is a sort of background murmur in the literature suggestion the existence of old affixes with semantic content which derive from frozen classifiers (e.g. Costello 1996, 1998). For example, Thomas (1969: 105) gives evidence for a *sa-* prefix in Chrau which denotes animals. She says 'For the most part the first syllable is never dropped, except in direct address' which of course does mark its optionality in the minds of speakers. Although she was unable to find comparative evidence, in fact this prefix appears to quite widespread, as Shorto (2006: 469) notes a number of cognates. Table 1 cites cognates for the Chrau term for 'bear' [the animal] which shows that the s- prefix occurs in Bahnaric, Katuic and Vietic, with further possible cognates for the root itself in Aslian and Pearic.

Table 1. An Austroasiatic root for 'bear' with variable prefixes

Attestation	Language	Subgroup
kaw.ip	Jahai	Aslian
*c.kaw ~ *gaw	proto Bahnaric	Bahnaric
h.kaw	Laven [Jru']	Bahnaric
rə.kǫw	Sedang	Bahnaric
si.kaw	Chrau	Bahnaric
haŋ.kaw	Ngeq	Katuic
sa.kaw	Bru	Katuic
kəw.x̱aj su:ʔt	Chong [of Kanchanaburi])	Pearic
gấu	Vietnamese [Hanoi])	Vietic
cə.ku:	Chứt [Rục])	Vietic

Smith (1975) points to the widespread presence of a velar prefix for animal names in Sedang, and widely in the Vietnamese languages he sampled. However, this prefix is found across Austroasiatic and also, strikingly, widely in Sino-Tibetan. Matisoff (1973) draws attention to its presence in Lolo-Burmese, but as Benedict (1990) notes, it is present on the words for 'tiger' and 'hawk' throughout the phylum.

In some languages of the region, the use of the fricative to mark animal names is very marked. Table 2 gives an example from Western Miji, an only doubtfully Sino-Tibetan language spoken around Nafra in Arunachal Pradesh. The palatal fricative /ʃ/ is the most common prefix, but I am assuming s~ʦ are probably its allomorphs.

[4] An early version of these ideas was presented at the 40th Colloquium on African Languages and Linguistics, Leiden, 29th-31st August, 2011.

Table 2. Animal names in Nafra Miji

Gloss	ʃ	s/ʦ
animal		stɕɔ̌
horse	ʃgrɔ	
stallion	ʃgrɔ mbŭ	
mare	ʃgrɔ mnĕʔ	
colt	ʃgrɔ i	
sheep	ʃgθɔʔ	
goat	ʃprn	
dog	ʃazi	
barking deer		tstshũ
deer		tstsə
flying squirrel	ʃbiã	
leopard	ʃnmu	
monkey	ʃbŏ	
musk deer		tstsɲăw
pangolin, anteater	ʃgʤɔ	
wild cat	ʃgrĕ	
wild dog	ʃkʃə	
sparrow		slĭʔ
ant	ʃɲi	
fish sp. I		sθŭ
fish sp. II		sviaʔ
fish sp. III		sgiɔʔ

The neighbouring Hruso language also shows an S- prefix for animals, although it shares almost no lexical cognates with Miji, except probably 'ant' *ʃn*.

Photo 1. Hruso shaman, Thrizino

Table 3. Hruso animal names with S- prefix

Gloss	Hruso
wild animal	sm ʧi
dog	ʃʎuɔ
bear	stɜ̄ɔ
otter	sz̄ɛ
rat	ʒmɔ
ant	ʃn
caterpillar	ʃblu
flea	sgzə
bloodsucking fly	sdz̄m
cobra	ʒ̄tɔ̆
python	ʒ̄ʃaba
snake sp. I	ʒmə
frog I	ʃʤa
snail	svankɔ̆

This suggests strongly that what is transferred is the idea of the semantics of a prefix rather than lexical borrowing. The neighbouring Koro language, which is structurally very similar to Hruso and Miji, seems to show no trace of S- prefixes.

Miji and the related Bangru also have a very marked m- prefix related to body parts, both for humans and animals. Table 4 shows a comparative list of Western and Eastern Miji as well as Bangru. The Bangru citations are orthographic, and on comparative grounds, assume the vowel following the m- prefix is

epenthetic. Forms in square brackets are cited for completeness, where one branch has an m- prefix and the other lacks it.

Table 4. Miji and Bangru body parts with an m- prefix

Gloss	W. Miji	E. Miji	Bangru	Comment
arm, hand	(m)gĭ	(m)gĭ	m(e)gey	No Tibeto-Burman cognates
beard	mɔmyuʔ	mmɯʔ	m(a)maŋ	Widespread Tibeto-Burman root, though not with m- prefix
bone	mriaŋ	mriaŋ	mnii	Possibly cf. Northern Naga *raŋ
brain	mɲɔʔ	mɲɔʔ		No certain external cognates, though cf. Bodic, e.g. Tshangla *ɳok taŋ*
breast	mɲu	mnɯʔ		m- 'body part prefix' plus widespread Tibeto-Burman etymon **nu(w)*
chest	mθm kʸu	mɭoŋ kᵩʔ		The kV- element has widespread Tibeto-Burman cognates. Note Puroik *tə kɯ*
chin	mugudza	mgutɕă		No Tibeto-Burman cognates
ear	mʒɔʔ	mzɔʔ	m(i)bwa	No certain external cognates, but cf. Memba *namʤo*
eye	mmreʔ	mreʔ		No Tibeto-Burman cognates
face	mgmiaʔ	mkmiaʔ	m(e)kwii/mekuyi	Matisoff (2003) proposes *#s.myal* for PTB. The best cognates are in Maraic, e.g. Lakher *h.mia*, but the velar preceding the Miji stem is of unknown origin.
finger	mgi tso		m(e)gey ʧowa	cf. 'arm'
flesh	mzaʔ	mʒaʔ		#sa is widespread in Tibeto-Burman, but this may be coincidence
heart	luŋ, [θɔm vʸu]		mloŋ	#luŋ is widespread in Naga complex languages
kidney	mkbð		mpega	Neither root has a Tibeto-Burman cognate
liver	mtn		m(a)tayiŋ	cf. Chin roots such as Thado *tʰin*, and possibly proto-Tani **zin*.
lungs			mloŋ wasayi	cf. 'heart'
mouth	mugð		m(i)niŋ	STEDT relates gɔ to proposed PTB #ku(w) but this seems unlikely. Some Tani languages have apparent cognates, e.g. Apatani a.gũ but this is not apparently proto-Tani. The Bangru form has no obvious cognates.
navel	mʃmay			No Tibeto-Burman cognates. The ʃ- appears to be an earlier prefix.
neck	[dmuzõ]		m(i)niŋri	Scattered attestations in Kuki and Chin, e.g. Lushai #riŋ. Possibly related to much more widespread #luŋ
nose	[ɲubyuŋ]		m(i)niiko	Miji has ɲi 'blow nose'. Chin languages have common *niit* for 'blow nose'
rib			mpelowa	No Tibeto-Burman cognates.
shoulder	mfa		mpoʧ	Miji has very scattered Tibeto-Burman cognates, e.g. Chinbon pá, though forms with a back high vowel are widespread. No obvious cognates for Bangru.
stomach	mrŏ		mulgu	

Table 4. Miji and Bangru body parts with an m- prefix

Gloss	W. Miji	E. Miji	Bangru	Comment
thigh	mləʔ		mur^{h}	
throat	mryɔnza			
tooth	mtr̄		m(e)t^{h}u	
vein	mdtʔ			
wrist	gi mvθɛ			
Animals				
horn	mʃǯɔ̌		m(e)ws	
tail	mdmray		m(u)lwe	
hump	mkb^{y}u			
tusk	mt̂ŭ			
udder	mŋŭʔ			
fur, feather	mɔmyŭʔ			
wing	mktɕi			

There is limited comparative evidence for a Tibeto-Burman m- prefix, see for example Matisoff (2008:183) on **m-ley*~**m-li* for 'penis'.

Forrest (1962), in an article not often cited, points out that Lepcha (Rong) has the same *kV-* prefix for animals noted for Palaungic and Khmer. Rong also shares other animal prefixes with Austroasiatic, for example, *luk/lun-* for animals and plants corresponds to Khasian *lyŋ-* and the *sV-* prefix mentioned above, which is also attested in Khasian and Palaungic. Another prefix shared between Rong and Austroasiatic is the nominaliser which forms abstracts in Rong, *nun/num-*, also widely attested in Austroasiatic.

As an example of how the *kV-* affix appears synchronically, Table 5 shows a widespread root for 'buffalo' attested in most branches of Austroasiatic[5]. The term is borrowed into Austronesian and gives us the common English name carabao. The attestations in different languages provide an example of the complex build-up of prefixes that characterise this type of morphology. Shorto (2006) reconstructs **krpiʔ* for PMK[6], but the evidence seems to better support either a back or central vowel and a final palatal, thus the suggestion *k.r.*pu*.y. I have analysed each synchronic form as a combination of a root, plus segmental affixes, each separated by a full stop. The proposal for the leftwards movement of the final palatal to the interior of the root is shown with a raised ^{y}, thus *p^{y}u*. The front vowels arise from the final –y being incorporated into the stem. Sometimes this is merely lost and the back vowel is retained or lengthened. Whether the earliest form had a three consonant cluster in initial position is debatable; the original could have been **r.pu* as in Khmuic, which subsequently gained a k- prefix, as part of the widespread animal class in Austroasiatic identified by Smith (1975). Proto-Khmuic must have had something like **g.r.pu* to explain the synchronic forms. In Vietic, the b/p of the root was lost and r→l, generating *k.l.Vw* structures. The final nasal in Mon is mysterious unless it arose under the influence of the k- prefix.

Table 5. 'Buffalo' *#k.r.pu.y* **in Austroasiatic**

Phylum	Branch	Language	Attestation	Formula	Comment
Austroasiatic		PMK (Shorto)	*krpiʔ	k.r.pi.ʔ	
Austroasiatic		PAAS (RMB)		k.r.*pu*.y	
Austroasiatic	Monic	Mon	preaŋ ဓြေ	r.*p^{y}u*.ŋ	
Austroasiatic	Monic	Nyah Kur	chəlo̱w	k.r.(p)*u*	? < Vietic
Austroasiatic	Vietic	proto Vietic	*c-lu	k.r.(p)*u*	
Austroasiatic	Vietic	Thavung	khuay[1]	k.r.(p)*u*.y	
Austroasiatic	Vietic	Pong	klow	k.r.(p)*u*	

[5] Munda has *bɔŋtel* throughout, which may be the same root with the *–tel* an old compound. Mangic languages have *vɔ*, which again could well be cognate but a lack of morphology makes this speculative.

[6] I use proto-Mon-Khmer when citing previous literature, but in general this terminology should be discouraged, as perpetuating an outmoded classification (Sidwell & Blench 2011).

Phylum	Branch	Language	Attestation	Formula	Comment
Austroasiatic	Khmeric	Khmer	krəbɤy	k.r.*pu*.y	? < Stieng
Austroasiatic	Pearic	Pear	krəpaːw	k.r.*pu*.y	
Austroasiatic	Pearic	Chong	kapaːw^{A}	k.p^{y}u	
Austroasiatic	Bahnaric	PNB	*kapɔː	k.*pu*	
Austroasiatic	Bahnaric	Sedang	kopôu	k.*puu*	
Austroasiatic	Bahnaric	Tampuon	kəpəu	k.*p^{y}u*	
Austroasiatic	Bahnaric	Bahnar	kəpoː	k.*p^{y}u*	
Austroasiatic	Bahnaric	PSB	*g~rəpuː	k.r.*pu*	
Austroasiatic	Bahnaric	Mnong	rpu	r.*pu*	
Austroasiatic	Katuic	Proto-Katuic	*krpiiw	k.r.*p^{y}u*	
Austroasiatic	Katuic	Pacoh	kərbɤː	k.r.*pu*.w	alligator; dragon [!]
Austroasiatic	Katuic	Chatong	karpiiw	k.r.*p^{y}u*	
Austroasiatic	Khmuic	Sre	rəpu	r.*pu*	
Austroasiatic	Khmuic	Chrau	gəpuː	k.*pu*	
Austroasiatic	Khmuic	Biat	rpuː	r.*pu*	
Austroasiatic	Aslian	Kensiw	kɛˈpaw	k.*p^{y}u*	unless < Malay
Austroasiatic	Aslian	Temiar	kəɹbau	k.r.*p^{y}u*	< Malay

'Buffalo' indicates clearly the morphological path these nouns characteristically take; multiple affixing develops as prefixes, originally nominal classifiers with a semantic coherence differentially drop consonants leading to highly diverse outcomes. An Austronesian language such as Rhade, which today has *kbao*, probably originally had a longer, more characteristic Austronesian form, and has restructured it under the influence of its Austroasiatic neighbours. A Daic language such as Nung has a synchronic form *tú vai*, i.e. classifier plus stem, which may have been borrowed from a Vietnamese Austroasiatic language. This could also explain deviant Katuic forms such as Katu *tariiq*, which would originally have resembled Chatong *karpiiw*. The k- prefix became a suffix, the stem consonant p- was deleted and a now unproductive t- prefix was added, perhaps on the model of the Daic nominal classifiers.

Table 6 is a second illustration of the k- prefix for animals in SE Asian languages, showing a common root for 'tiger', which is attested across phylic boundaries.

Table 6. The *#kVla* root for 'tiger' in SE Asian languages

Phylum	Branch	Language	Attestation
Sino-Tibetan	Sinitic	OCM	*hlâʔ
Sino-Tibetan	Burmic	Old Burmese	klya
Sino-Tibetan	Bodish	Monpa	khai-la
Austroasiatic	Khmeric	Angkorian Khmer	khlaa
Austroasiatic	Pearic	Samre	kanɔh^{A}
Austroasiatic	Bahnaric	PB	*kəlaa
Austroasiatic	Bahnaric	Sedang	klá
Austroasiatic	Katuic	Pacoh	kulaa
Austroasiatic	Katuic	Ir	kalaʔ
Austroasiatic	Katuic	So	kula
Austroasiatic	Palaungic	Shinman	kaʔ^{4} vai^{3}
Austroasiatic	Monic	Proto-Monic	*klaaʔ
Austroasiatic	Aslian	Sakai	kla
Austroasiatic	Khasian	War Jaintia	k^{h}la
Austroasiatic	Muṇḍā	Muṇḍā	kula
Daic	Tai	Thai	kla

The root for 'tiger' is probably an illustration of how semantically assigned affixes are borrowed. The tiger is an animal of great symbolic importance across the region and has probably been borrowed extensively,

fossil morphology and all. Once a semantic association of a k- prefix for 'animal' is set up (cf. Table 5 above) it is easily generalised to other animals within a particular speech.

3. Word structure in Sino-Tibetan

The Sino-Tibetan language phylum has a disputed membership, internal structure and thus debatable reconstructed forms. Van Driem (2008), Handel (2008) and Blench (2008) give an overview of some of the key issues. These swirl around the position of Sinitic, formerly considered a primary branching, but now often treated as simply another branch within Sino-Tibetan, hence its rechristening in some authors as Tibeto-Burman. Proposals to rename the phylum in a more neutral fashion (e.g. Tibeto-Burman or Trans-Himalayan) certainly have merit. It is certainly the case that Sinitic shares far more lexically with common Tibeto-Burman lexicon than some of the isolated groups of Arunachal Pradesh (Blench & Post in press). As with Austroasiatic, Sino-Tibetan words tend to have a core segmental structure, very often CV(N), and then one or more affixes, both prefixed and suffixed. Affixes can shift right or left into root medial position, inducing consonant and vowel changes, and prosodies affecting C_1. Change in the segmental character of C_1, such as n→ɲ or ŋ, is deemed to be driven by a shift of palatalisation or velarisation into the core. The perceived incorporation of a consonant within the stem leads to affix renewal, and thus stacking of unproductive morphemes. This is also very common in phyla such as Nilo-Saharan, where languages like Krongo (Kadu group) have up to three unproductive affixes preceding nouns (Reh 1985).

Table 7 is intended to demonstrate how this works in Tibeto-Burman with the root for 'two', not strictly a noun, but subject to analogous processes. The starred forms are drawn from standard sources, and are not necessarily endorsed, merely cited for a convenient comparison. It is assumed there was a core *ni*, with a velar prefix and two suffixes, a fricative and a high front vowel or an approximant. The velar prefix was regularly suffixed and weakened to ʔ. The fricative suffix was either affricated or weakened to –h and switched to a prefix. Forms like Cho *hngih* may represent copying, leaving the affix appears at both ends of the word. Other more sporadic affixes are added, such as p-, t-, r- and possibly a-.

Table 7. The root C.ni.C(C) for 'two' in Tibeto-Burman

Language	Group	Form	Formula
*Sino-Tibetan	Sino-Tibetan	gnyis	g.nʸi.s
*Tibeto-Burman	Tibeto-Burman	g-ni-s	g.ni.s
*Karen	Karenic	hni	h.ni
*Lolo-Burmese	Lolo-Burmese	ʔnit	ʔ.ni.t
*Loloish	Loloish	s-ni(k)[2]	s.ni.k
*Northern Naga	Northern Naga	ʔ-ni	ʔ.ni
Bugun	Bugun	ɲeŋ	nʸi.ŋ
Taraon	Mishmic	kaiŋ	k.ni
Idu	Mishmic	kaɲi	k.nʸi
Puroik	Puroik	ɲi	nʸi
Kamengic	Mey of Shergaon	ɲit	nʸi.t
Miji	Miji	gni	g.ni
Miju	Mijuish	knîn	k.ni.n
Koro	Siangic	ki-ne	k.ni
Milang	Siangic	nə	ni
Karbi	Mikir	hiní	h.ni
Meithei	Meithei	ə-nì	ə.ni
Newar (Dolakhali)	Newar	nis	ni.s
Atong	Bodo-Garo	ni	ni
Garo	Bodo-Garo	gəni	g.ni
Kokborok	Bodo-Garo	nəy	ni.y
Ao (Mongsen)	Naga	anət	a.ni.t
Rongmei	Naga	kənə̃i	k.ni.y
Tangkhul	Naga	[3]khə [3]ni	k.ni
Phom	Northern Naga	ñi[31]	nʸi
Cho (Mindat)	Chin	hngih	h.n(g)i.h

Language	Group	Form	Formula
Daai	Chin	ŋ̣ɲ̣iʔ	ŋ.ni.ʔ
Khumi	Chin	n̥ue(ng)	ni.ŋ
Lai (Hakha)	Chin	pa-hniʔ	p.h.ni.ʔ
Lakher [Mara]	Chin	³sa ²nɒ	s.ni
Lakher [Mara]	Chin	pā-nō	p.nʷi
Lushai [Mizo]	Chin	hnih	h.ni.h
Matu	Chin	pa̰ɴiʔ	p.ni
Nyhmoye	Chin	ŋ̣ɲ̣iʔ	ŋ.ni
Bhramu	Himalayish, Western	n̥is	ni.s
Kanauri	Himalayish, Western	nis	ni.s
Motuo Menba	Monpa	ŋ̊ik tsiŋ	nʸi.k
Kaike	Bodic	nghyi	g.h.nʸi
Tshona (Mama)	Bodic	nʌi¹³	a.ni
Tibetan (Alike)	Tibetic	ɣŋ̊i	g.nʸi
Tibetan (Amdo: Bla-brang)	Tibetic	hŋ̊i	h.nʸi
Tibetan (Balti)	Tibetic	ŋis	n(g)i.s
Tibetan (Sherpa)	Tibetic	ngyi	g.nʸi
Dirang	Tibetic	nitsiŋ	ni.ts.ŋ
Tawang	Tibetic	neⁱ	ni
Memba	Tibetic	ɲi	nʸi
Meyor	Tibetic	ni	ni
Burmese (Written)	Burmish	hnaʦ	h.ni.ʦ
Marma	Burmish	hnɔiʔ	h.nʷi.ʔ
PNL	Loloish	ʔnitᴸ	ʔ.ni.t
Ahi	Loloish, Central	ni²¹	ni
Lalo	Loloish, Northern	ni²¹	ni
Nasu	Loloish, Northern	ŋ̊i⁵⁵	nʸi
Nusu (Southern)	Loloish, Northern	ɦĩ̃³⁵	h.ni
Akha	Loloish, Southern	nyì	nʸi
Mpi	Loloish, Southern	ɲiʔ²	nʸi.ʔ
Naxi	Naxi	ŋ̊i²¹	nʸi
Chinese (Old)	Sinitic	njijs	nʸi.(y)s
Sak	Luish	níŋ-hvú	ni.ŋ
Anong	Nungic	əni	ə.ni
Dulong	Nungic	ɑ³¹ ni⁵⁵	a.ni
Nung	Nungic	ɑ³¹ ŋ̊⁵⁵	a.nʸi
Ersu	Qiangic	nɛ⁵⁵	ni
Guiqiong	Qiangic	ŋ̊i³³	nʸi
Namuyi	Qiangic	ŋ̊i⁵³	nʸi
Qiang (Mawo)	Qiangic	ɣnə	g.ni
Tangut [Xixia]	Qiangic	njɨ̲	nʸi
Caodeng	rGyalrongic	ʁnes	r.ni.s
Daofu	rGyalrongic	ɣnə	g.ni
rGyalrong	rGyalrongic	kěněs	k.ni.s
Gurung (Ghachok)	Tamangic	ŋ̃ĩhq	ŋ.nʸi.h.q
Tamang (Sahu)	Tamangic	'nyi:h	nʸi.h
Thakali	Tamangic	'ngih	n(g)i.h
Kayan (Pekon)	Karenic	θaní	t.ni
Magar	Kham-Magar	nis	ni.s
Thulung	Kiranti	nək	ni.k
Limbu	Kiranti, Eastern	nɛccʰi	ni.s
Bahing	Kiranti, Western	nik-si	ni.k.s
Apatani	Tani	tá-ñe	t.nʸi

Language	Group	Form	Formula
Nah	Tani	a-ɲi	a.nʸi
Tujia	Tujia	ɳ̊ie^{55}	nʸi.V

The rapid switching and replacement of affixes in Tibeto-Burman shows all too clearly the defects in the usual process of reconstruction. The 'method', such as it is, involves choosing a common segmental core and then proposing the most commonly attested affixes to accompany it. But common affixes may well be evidence for lower-level nodes, or indeed diffusion. The similarities between affixes attested in Austroasiatic and Sino-Tibetan show that these can spread both from language to language and across phylum boundaries.

Sinitic historical phonology allows us to see these processes as they occur. Early Zhou Chinese has many more affixes familiar from other Tibeto-Burman languages than its later descendants. Table 8 shows a set of lexemes attested in Zhou which are found either with fewer or without affixes in later forms.

Table 8. Affix movement and loss in the evolution of Chinese

Gloss	Early Zhou	Classical	Character
blood	s.wiːt	wiːt	血
fire	s.məːyʔ	m̥ǎːy	火
head	s.luʔ	l̥ǔ	首
black	s.məːk	m̥əːk	黑
see	keːn.s	keːn.s	見
seed	toŋ.ʔ	tǒŋ	種

To illustrate the convergence of Sino-Tibetan and Austroasiatic affixes in semantics, Table 9 shows one of the principal roots for 'bear' in Sino-Tibetan. The original form may have been something like *twŏ.m* , currently attested in rGyalrong. This would account for many synchronic forms with roots such as *vom, wom, hom*, with or without affixes. However, strikingly, the common prefix for 'bear', as for some other animal names in Sino-Tibetan, is s-, just as in Austroasiatic (§2.).

Table 9. 'Bear' in Tibeto-Burman languages

Group	Language	Form	Formula
Central Loloish	Kucong	sa^{35} mu^{31}	s.-m
Chin	Lakher [Mara]	chā-vỳ	s.vo.m
Chin	Lushai [Mizo]	sà-váwm	s.vo.m
Kham-Magar-Chepang-Sunwar	Chepang	siŋʔ.tyamh.yom	s.t.vo.m
Meithei	Meithei	shaum	s.wom
Naga	Lotha Naga	sēvá̱n̄	s.vo.m
Naga	Ao (Chungli)	shim	s.hʸom
Naga	Khoirao	chawom	ʧ.wom
Naga	Lotha Naga	seva	s.vo(m)
Naga	Maram	sahom	s.hom
Naga	Rongmei	cagüm	ʧ.g.wom
Naga	Tangkhul	¹si ¹ŋom	s.g.wom
Nungic	Rawang	ʃəwi^{53}	s.wʸo(m)
Sinitic	Chinese (Old/Mid)	g̱ium/j̱iung	g.yom
Loloish, Southern	Akha	xhà-hḿ	g.hom
Loloish, Southern	Hani (Khatu)	sjhí	ʃ.hʸo(m)
Tani	Galo	sotum	s.tom
Tani	Padam-Mising [Abor-Miri]	si-tum	s.tom
Tani	Apatani	si-tĩ	s.tʸo
Tani	Bengni	šu-tum	s.tom
Tani	Bokar	šu-tum	s.tom

This is not the place to develop a specific historical model to explain this situation. But what this can be taken to demonstrate is;

a) Sino-Tibetan and Austroasiatic have underlyingly similar word structures, without being genetically related
b) that the so-called 'minor syllable' is an optional affix, which often has semantic content, and which can be shifted to a suffix, or incorporated into the stem
c) that unproductive affixes can be subject to renewal, for example reprefixing, without forming consonant clusters
d) that semantic associations of affixes both can be borrowed across phylic boundaries, along with the segmental material, and evidently were at an early stage of the evolution of these phyla

4. Contact, borrowing and metatypy

The Austroasiatic and Sino-Tibetan language phyla are intertwined across much of their geographical range today and we should expect considerable local borrowing. The Munda languages are cut off from the remainder of Austroasiatic by a zone of highly diverse Sino-Tibetan languages. There are evidently two distinct issues, local borrowing and broader structural similarities between the two phyla. Studies of this issue are sparse; Benedict (1990) discusses Austroasiatic loans in Sino-Tibetan and Shafer (1952) is a study of similarities between Khasi and Sino-Tibetan, evaluated in Diffloth (2008). Forrest (1962) and Bodman (1988) both discuss the puzzling issue of apparent Austroasiatic similarities in Lepcha (Rong), a language no longer in direct contact with Austroasiatic.

Although there *are* deep-level lexical borrowings between Austroasiatic and Sino-Tibetan, they appear to be few (Benedict 1990). Most striking, however, are the similarities of word-structure and affixes, often displaying the same segmental material with comparable semantics. What seems to have occurred is extensive metatypy, i.e. long-term bilingualism causing convergence of structures. The infrequency of lexical borrowing must be due to sociolinguistic factors, a desire for esoterogeny, marking the separateness of languages. This is probably at its most extreme in Arunachal Pradesh, where neighbouring languages with extremely similar cultural concepts, such as Miji, Hruso and Koro, share no more lexical cognates than could be expected by chance.

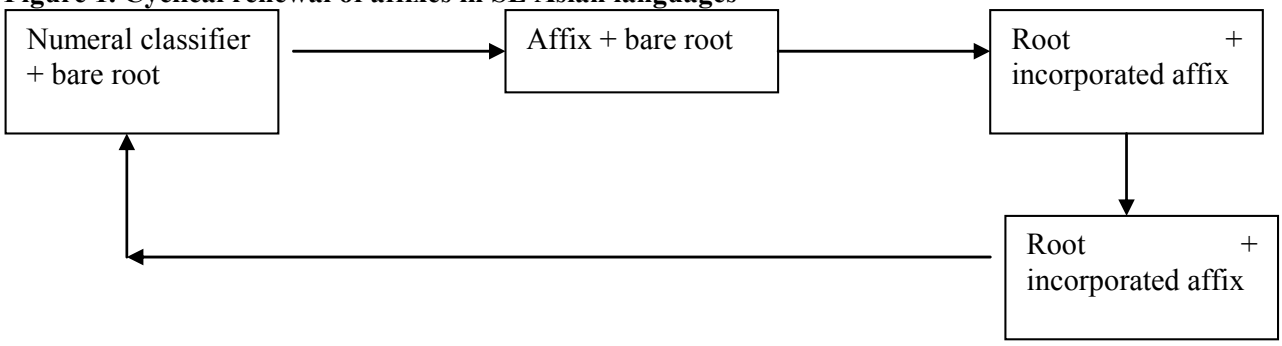
The key to these convergent structures is the incorporation and re-analysis of numeral classifiers. In many Eurasian and New World language phyla, nominal plurals are marked with short morphemes that indicate to hearers the semantic class of object under discussion. Sino-Tibetan and Austroasiatic, along with many other language phyla (Austronesian and many New World phyla) are sporadically marked by nominal classifier systems. Indo-European languages only have this system to a very limited extent; a 'flock of sheep', a 'herd of cattle', a 'bunch of bananas'. Classifiers are essentially grammaticalised nouns that have become obligatory accompaniments (clitics?) when marking plurals or groups of nouns (Aikhenvald 2000). Nominal (or numeral) classifiers are a distinctive feature of SE Asian languages widespread in Austroasiatic (Adams 1989), Sino-Tibetan (e.g. Post 2007), Austronesian and Daic (Jones 1970). There has been much discussion about the difference between classifiers and quantifiers, but essentially, classifiers are non-concordial lexemes that are obligatory with certain classes of plural nouns. In some languages they can act as pronominals, for example in Chrau (Thomas 1969).

There is strong evidence for borrowing between phyla; for example, one of the 'animal' classifiers in Austroasiatic is probably a borrowing from Daic. Classifiers often have a transparent etymological origin, typically grammaticalised body parts (see e.g. Becker 1975 for Burmese), and this process of grammaticalisation has been similarly documented for noun-class affixes in African languages.

Both Austroasiatic and Sino-Tibetan may originally have had simple stems, with no affixes marking number, case, semantics or gender. Numeral classifiers, usually CV(C) syllables with semantic assignations, and were put together with nouns, usually preceding them, as is still very much the situation in Daic languages. It is possible that Sino-Tibetan and Austroasiatic numeral classifiers became bound to the root and reduced to C with an epenthetic vowel following, hence the transformation into affixes. Although this occurred to a

greater or lesser extent in different languages, consciousness of their separateness was retained. As a consequence, they can be shifted to the end of the root, deleted in some languages and a new prefix added, or a new prefix added preceding the existing prefix. Meanwhile, distinct numeral classifiers continued to co-exist and continued to be incorporated and renewed. Figure 1 shows a highly schematic visualisation of this process of renewal; the examples in the text indicate some of the complexities encountered on the way.

Figure 1. Cyclical renewal of affixes in SE Asian languages



The analytic problem with this hypothesis is that classifiers are usually very numerous and diverse. The lists given in Adams (1989) or Post (2007) are extremely long, and are moreover, strongly concerned with shape and appearance, something characteristic of Bantu, but not necessarily marked C affixes. I can offer one piece of evidence which is suggestive, the *m-* prefix in Mey [=Sherdukpen] of Rupa and the widespread *ma(ŋ)(k)* classifier for fruits, found in Tai languages.

Table 10. m- prefix for fruits in Mey of Rupa

Gloss	Rupa
fruit	m.laŋ
banana	m.suŋ
lemon	m.kẽ
sugar-cane	m.ʧĩ
walnut	m.ku

If we take the example of the mangosteen, *Garcinia mangostana*, the name of this fruit originates in Thai, with the classifier maŋ preceding it. The term is borrowed into Lao (where it remains meaningful within the classifier system) and Khmer and Burmese, where the Thai classifier makes no sense (Blench 2008). Table 11 shows how the word was borrowed across language phylum boundaries in MSEA languages.

Table 11. 'Mangosteen' in MSEA languages

Phylum	Language	Transcription	Script
Daic	Thai	máŋ kʰút	มังคุด
Daic	Lao	máŋ kʰūt	ມັງຄຸດ
Sino-Tibetan	Burmese	mĩ́ guʔ	မင်းကွတ်
Austroasiatic	Khmer	mʊəŋ kʰut	មង្គុត
Austroasiatic	Vietnamese	măŋ³³ kut²¹	măng cụt

By what process of copying and re-interpretation the *mVn* classifier eventually reached Mey will be difficult to recover. But I offer this as an example of the sort of process responsible for the semantics of affixes in Sino-Tibetan and Austroasiatic.

Photo 2. Mangosteen

Most languages in the region also historically had suffixes; it seems likely these were also originally classifiers and indeed, the similarities of segmental material suggest that prefixes became suffixes. Although it has been suggested by typologists that such a process is rare, Harvey

et al. (2006) reference some of the literature while illustrating this for Northern Australia (see also Green 1995). The evolution of the noun-class suffixes in Gur and Adamawa languages within Niger-Congo is similarly an example of prefix-suffix shift, carrying all the segmental material and concomitant alliterative concord. Within SE Asia, suffixes tend to weaken to glottal stops, nasals or approximants. As the final segments of a word erode, their features are incorporated into the stem, resulting in unstable vowels and vowel length. This is very similar to the sort of word shortening characteristic of languages of the Cameroun Grassfields, where final syllable erosion is responsible for complex tones (e.g. Mambiloid).

5. Parallel processes in African languages

Apart from Afroasiatic, African languages are usually considered to characterised by noun-classes, and not to have nominal classifiers at any historical time-depth. However, there are some interesting cases where these appear to be developing, for example Kana, a Cross River language in the Niger Delta of Nigeria (Ikoro 1996). The relatives of Kana are classic nominal affix alternation languages, but Kana seems to have reduced this system but compensated by developing nominal classifiers through grammaticalisation (Williamson 1985). More surprising is the case of Mambay, an Adamawa language of northern Cameroun, which still has a quite prominent concordial nominal suffix system, but which is developing prefixed nominal classifiers (Anonby 2011). Examples given by Anonby include the 'collectives' which precede nouns, do not show concord and have broad semantic correlations.

Nilo-Saharan is more puzzling, since its nominal morphology is marked by extensive affix alternation, and yet there is no system of alliterative concord and no systematic association with semantic categories. However, evidence seems to be increasingly emerging that we have been quite wrong in our understanding of Nilo-Saharan, that its underlying morphology is a system of nominal classifiers. Various studies have noted associations between affixes and semantic themes in different branches. For example, both Stevenson (1991) and Gilley (in press) note the semantic associations of affix pairings in Kadu languages, and Storch (2005) analyses these for Western Nilotic. Carlin (1993) observes some So number markers have broad semantic themes. But the most striking evidence comes Gumuz, a Nilo-Saharan language of the Ethio-Sudan borderland, whose Mayu dialect has been studied by Ahland (2010). Gumuz has a system of nominal incorporation, where a series of body part nouns are incorporated into verbs and which 'classify' the object, or more rarely the subject or instrument.

Gumuz turns out to have a system of predicate classifiers, marking semantic fields, typically of shape or texture. These are infixed in 'split verbs' and are copied as demonstratives. The major classifiers are -Vk'^{w} 'head', -Vts 'body', -Vc 'eye/seed', -k'^{w}ós 'tooth', and –ts'ê 'ear'. Mithun (1986) describes a verbal classifier whereby "a noun is incorporated into a verb to categorize an extra predicate argument...usually in S or O function." With this type of verbal classifier, there is frequently a generic-specific relationship between the incorporated NP and the external NP which accompanies it. The significance of this system is that classifiers which develop from grammaticalised body parts are governed by the semantics of nouns. For example; *'entities that are head-like in shape and/or function or closely associated with such objects'* govern the following classes of object 'fingers, toes water, sauce, beer, lotion, soap (in a container) ears of corn pots, pans, cans'. In constructions where the classifier refers to the object of the main verb, the classifier is suffixed to the verb and thus abuts the object noun directly. It is thus not difficult to see how it could become attached to the noun rather than the verb. Western Nilotic, as described by Storch (2005), has a system of nominal suffixes which appear to have semantic correlates There are both singular and plural suffixes marking length, roundness, part of etc. And there is some evidence for a system of suffix alternation which has been overwritten by the diffusion of 'imperial' number markers kV- and N- from other branches of Nilotic. As with Gumuz, grammaticalisation of body parts seems to be a major source for affixes. The T/K and N/K 'substrata' identified by Bryan in the 1950s and Greenberg's (1981) 'moveable –k as a Stage III article' are all reflections of this broader phenomenon.

Exactly how the Niger-Congo and Nilo-Saharan systems are related is still under discussion. Some branches of Niger-Congo, such as Mande, Dogon, Ijoid and Kaalak-Domurik, show no clear traces of any affix system. However, the remainder, Atlantic, Kwa, Benue-Congo and Gur-Adamawa have elaborate noun-classes and alliterative concord, or traces of such systems where they have demonstrably been lost (e.g. in Volta-Niger and Kru). So this system develops *within* Niger-Congo (and is thus probably not to be

reconstructed to proto-Niger-Congo, despite an extensive literature to the contrary). Other evidence, such as labial-velars, ±ATR vowels etc. point to extensive contact between Nilo-Saharan and Niger-Congo, and it is not stretching credibility to propose that the noun-classes of Niger-Congo represent a regularised metatypy of Nilo-Saharan affix systems. In other words, something that was implicit in the Nilo-Saharan system of nominal classifiers was borrowed as a system into part of Niger-Congo and then analogised as a rich system of nominal classes[7].

From this perspective, the similarities with SE Asian languages become clearer. Affix renewal is an extremely common process in Niger-Congo, where noun-class affixes become unproductive and a new affix is added (e.g. Childs 1983). Usually, however, the unproductive affix retains its vowel, or else the conjunction of two consonants results in a complex consonant. For example, Hyamic (Plateau) has developed a complex system of alternating initial clusters due to deletion of –V in the prefix. In SE Asia, the weakening of -V does not result in consonant merger but is rather retained as a syllable sequence. Such a word structure is not typical of Niger-Congo in general, but at least one group of languages does look like this synchronically. Nouns in the northwest Kainji languages (cLela, tHun, ut-Main, Gwamhi-Wuri), typically have the structure C.CVCV (the prefix is often transcribed with a schwa to make sense of the otherwise disquieting appearance) (e.g. Hoffmann 1968). Table 12, shortened from Paterson (2012), shows the noun-class prefixes of Ut-Ma'in, some of which now only have consonant prefixes, but which retain strong semantic associations. ə̄ marks the epenthetic vowel for single-consonant prefixes.

Table 12. Ut-Ma'in noun class prefixes

Class	sg.	pl.	Object Pronoun	Examples
1u	ū-	ú/wá	ū-mákt	'barren woman'
1ø	ø-	wá	ø-hámə̀t	'visitor'
2	ø-	ɛ́	ø-ná	'oxen, bovines'
3u	ū-	ɔ́	ū-bù	'house'
3ø	ø-	ɔ́	ø-bòʔ	'dream'
4	ə̄s-	sɛ́	ə̄s-bòʔ	'dreams'
5	ə̄r-	dɛ́	ə̄r-kɔ́k	'calabash'
6	ə̄t-	tɔ́	ə̄t-kɔ́k	'calabashes'
6m	ə̄m-	mɔ́	ə̄m-nɔ̀ːg	'oil'
7u	ū-	já	ū-ná	'ox, bovine'
7ø	ø-	já	ø-tʃə̄mpá	'man'
AUG	ā-	á	ā-kɔ́k	'huge calabashes
DIM	ī-	ɛ́	ī-kɔ́k	'tiny calabash'

These prefixes can be said to bear tone, although as it appears to be always low, it is no longer functional (also the case in Himalayan Sino-Tibetan). Similarly, many Kordofanian[8] languages have C.VCV structures, where the initial C is an alternating prefix. Schadeberg (1981a, 1981b) illustrates this for both the Heiban and Talodi groups. For example, Table 13 shows the reconstructed noun-class prefixes of Proto-Heiban (Schaeberg 1981a: 133).

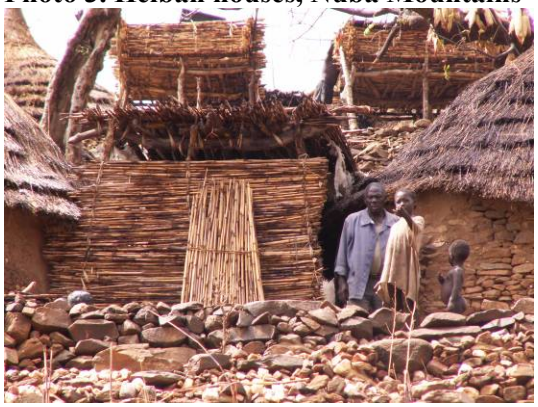
[7] The origins of alliterative concord can be debated, but a stimulating suggestion is the proposal of Hoffmann (1968) that demonstratives which copy affixes can explain the movement from prefix to suffix. Extending this idea, if affixes became re-analysed as separable, they can easily become demonstratives or articles, and once copied, establish the principle of alliteration.

[8] 'Kordofanian' is a creation of Greenberg (1963) based on the assumption that the Niger-Congo languages of the Nuba Mountains must form a genetic group, although this now looks like an over-optimistic view. However, these languages do share common morphological features, which is perhaps due to contact.

Table 13. Proto-Heiban noun-class prefixes

sg.	pl.
gu-	li-
g-	j-
d-	n-
li-	bu-
ŋ-	ɲ-
d̠-	d-

Photo 3. Heiban houses, Nuba Mountains



The difference with northwest Kainji is that the typical Kordofanian stem is -CVC. This suggests (perhaps) loss of C_1 of the stem, subsequent loss of –V from the prefix or assimilation of the resultant VV sequence.

The overall parallels to be drawn with African languages are as follows;

a) Nilo-Saharan languages have traces of a former numeral classifier system, still realised in Gumuz, which surface synchronically as moveable affixes and which have sporadic semantic associations
b) A subset of Niger-Congo languages have noun-class affixes with semantic associations, although these are regularly lost and re-evolve
c) These affixes can be shifted, disappear, fossilise or be incorporated into stems, leading to a process of renewal.
d) These affixes typically conserve their co-associated vowel, because it has a strongly associated segmental tone, whereas SE Asian languages weaken the vowel because there is no underlying tone.
e) However, Niger-Congo languages can occasionally lose the -V- of the affix so comprehensively that the result is a segmental affix consisting only of C, with resultant structural similarities to SE Asia

For these reasons, it seems that outcomes in SE Asia, while diverse, are still less exuberant than in West Africa. Excluding Bantoid and Bantu, the principal branches of Benue-Congo are Kainji, Plateau, Jukunoid, Cross River and ? Ikann. All of these have evidence (and usually synchronic examples) for a Bantu-like system of alternating nominal prefixes marking and alliterative concord applied to adjectives and other parts of speech. From this it is reasonable to conclude that nouns in the system of the proto-language had a basic (C)V.CVCV morphology, assumed by De Wolf (1971) in his now outdated study[9]. If so, the surface forms we see today are a development from this. However, those surface forms are astonishingly diverse. As an example of the complexity with a single family, Table 14 illustrates the possible outcomes from this type of restructuring.

[9] De Wolf reached this view by working 'backwards' from Bantu rather than actually surveying Benue-Congo as the primary data source.

Table 14. Erosion and restructuring of CV nominal prefixes within Benue-Congo

Strategy	Languages
Complete loss of affixes	Shen
Complete loss of prefixes, vowel neutralisation in stem, loss of C_2, number marked by contrastive vowel length	Cara
Complete loss of prefixes, addition of generalised prefix, suffixed or clitic plural marker	Lower Cross, Central Jukunoid
Complete loss of prefixes, development of nominal classifiers	Kana
Complete loss of original affixes, new affixes added on the basis of reduplication of first syllable of stem	Hasha
Loss of V of affixes leading to C- prefixes	Northwest Kainji (cLela etc.)
Loss of V of affixes leading to long C- stem initials	Kambari, Upper Cross, Jju cluster
Loss of V of affixes leading to stem initial consonant clusters and consonant alternation	Hyamic
Existing affixes become frozen to the stem and are reprefixed	Cibər [Lopa]
Reduction of all CV- prefixes to V-	Ikann, some Plateau
Reduction of all CV- prefixes to u/i- and rightwards shift into stem, leading to contrastive palatalisation and labialisation	Many Plateau
Reduction of RV- and NV- prefixes to R-, N-, and rightwards shift into stem, leading to sporadic nasalisation and rhotacisation	Many Plateau
Prefixes become suffixes	Some Mambiloid
Prefixes become suffixes, which are deleted producing complex stem tones	Mambila
Prefixes become suffixes, lose final –V, C is frozen to the stem and number marking is lost	Dakoid
Prefixes partly become suffixes, resulting in systems of double-affixing	Tivoid

Once the descriptive language is changed, many of these processes are also attested in Austroasiatic and Sino-Tibetan.

6. A historical scenario

Describing structural similarities is one thing, accounting for them historically is quite another. A neat explanation would have Sino-Tibetan and Austroasiatic originating in neighbouring areas and these similarities be phenomena deriving from early contact. However, this is difficult to support using current hypotheses about homelands. Austroasiatic has a long history of quite varied speculations about its homeland (see review in Sidwell & Blench 2011). For scholars who still support the Mon-Khmer hypothesis, the Bay of Bengal is an option. Diffloth (2005) has generally argued for a southern, tropical locus on the basis of faunal reconstructions. Sidwell & Blench (2011) propose a riverine dispersal from the Central Mekong, based on their parallel array model of Austroasiatic classification. Sino-Tibetan has a similarly varied menu of hypotheses, from the views of Matisoff ('the flanks of the Himalayas'), Van Driem (1998) Sichuan and Blench & Post (in press) arguing for Northeast India.

Unless these hypotheses are very misguided, proximate homelands are not the solution. Austroasiatic clearly spread far and fast, probably along the river systems of SE Asia, seeking humid valleys to grow taro while using improved boat technology. Only such a hypothesis would account for the arrival and diversification of the Munda languages in India. If the proposals in Sidwell & Blench (2011) are correct, then this would have been around four thousand years ago, when there is a rapid and sudden expansion of the Neolithic in mainland SE Asia, marked by the spread of 'incised and impressed' pottery (Rispoli 2009). So there may have been intensive contact between Austroasiatic and Sino-Tibetan in the zone between northern Vietnam, Laos and northeast Myanmar, and consequent diffusion of key structural traits. Purely chronologically, these traits are likely to originate in Sino-Tibetan, as they are clearly attested in Sinitic and many of the highly diverse languages of NE India. Probably this question cannot be fully resolved until we have better mapping of the distribution of semantically significant prefixes across multiple language phyla.

7. Conclusions

It is unlikely that SE Asian specialists will find this analysis very palatable; the established terminology works hard against the notion of noun classes, a feature usually associated with Africa, Papuan and Australian, as well as some Amazonian languages. But research traditions and proposals for reconstruction also do not seem very credible, if proto-forms simply pile up unexplained affixes. The next step in the linguistic prehistory of the region is exploring its historical morphology in much greater detail and in particular accounting for the remarkable structural convergence at a particular historical juncture.

References

Adams, Karen Lee 1989. *Systems of numeral classification in the Mon-Khmer, Nicobarese and Aslian subfamilies of Austroasiatic*. Canberra: Pacific Linguistics.

Ahland, Colleen A. 2010. Noun incorporation and predicate classifiers in Gumuz. *Journal of African Languages and Linguistics*, 31(2): 159–203.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Anderson, Gregory D.S. 2004. Advances in Proto-Munda reconstruction. *Mon-Khmer Studies* 34.159–184.

Anonby, Erik 2011. *A grammar of Mambay*. Köln: Rüdiger Köppe.

Becker, Alton L. 1975. A linguistic image of nature: the Burmese numeral classifier system. *International Journal of the Sociology of Language,* 5: 109-121.

Benedict, P. 1990. Austroasiatic loanwords in Sino-Tibetan. *Mon-Khmer Studies Journal*,18-19: 1-13.

Blench, Roger M. 1995. Is Niger-Congo simply a branch of Nilo-Saharan? In: *Proceedings of the Fifth Nilo-Saharan Linguistics Colloquium, Nice, 1992*. ed. R. Nicolai and F. Rottland. 68-118. Köln: Rudiger Köppe.

Blench, Roger M. 2006. *The Kadu languages and their affiliation: between Nilo-Saharan, Niger-Congo and Afro-Asiatic*. Insights into Nilo-Saharan Language, History and Culture. Al-Amin Abu-Manga, L. Gilley & A. Storch eds. 101-127. Köln: Rüdiger Köppe.

Blench, Roger M. & M. Post (in press) Rethinking Sino-Tibetan phylogeny from the perspective of North East Indian languages. Paper accepted for a volume of selected papers from the *16th Himalayan Languages Symposium* 2-5 September 2010 School of Oriental and African Studies, London. Nathan Hill & Tom Owen-Smith eds. Mouton de Gruyter.

Bodman, Nicholas C. 1988. On the place of Lepcha in Sino-Tibetan, a lexical comparison. *Linguistics of the Tibeto-Burman Area*, 11(1): 1-26.

Carlin, Eithne 1993. *The So language*. AMO2. Köln: Universität zu Köln.

Childs**,** Tucker 1983. Noun class affix renewal in Southern West Atlantic. In: *Current Approaches to African Linguistics Volume 2*, eds. J. Kaye et al. 17-29. Dordrecht: Foris Publications.

Costello, N. 1966. Affixes in Katu. *Mon-Khmer Studies Journal*, 2: 63-86.

Costello, N. 1998. Affixes in Katu of the Lao P.D.R. *Mon-Khmer Studies Journal*, 28: 31-42.

De Wolf, P. 1971. *The noun class system of Proto-Benue-Congo*. The Hague: Mouton.

Diffloth, Gérard 2005. The contribution of linguistic palaeontology and Austroasiatic. In: Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics.* 77-80. London: Routledge Curzon.

Diffloth, Gérard 2008. Shafer's parallels between Khasi and Sino-Tibetan. In: *Northeast Indian Linguistics*. S. Morey and M. Post eds. 93-104. New Delhi: Cambridge University Press.

Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42(2): 214-261.

Enfield, N.J. 2007. *A grammar of Lao*. Berlin/New York: De Gruyter.

Forrest, R.A.D. 1962. The Linguistic Position of Rong (Lepcha). *Journal of the American Oriental Society*, 82(3):331-335.

Green, Ian 1995. The Death of 'Prefixing': Contact Induced Typological Change in Northern Australia. *BLS*, 21:414-425.

Greenberg Joseph 1977. Niger-Congo noun class markers: prefixes, suffixes, both or neither. *Studies in African Linguistics*. Supplement 7, 97-104.

Greenberg Joseph 1981. Nilo-Saharan Moveable- k as a Stage III Article (with a Penutian Typological Parallel). Journal of African Languages and Linguistics, 3 (2): 105-112.

Greenberg, Joseph H. 1963. *Languages of Africa* (Indiana University Research Centre in Anthropology, Folklore, and Linguistics, Publication 25). The Hague: Mouton & Co.

Handel, Zev 2008. What is Sino-Tibetan? Snapshot of a Field and a Language Family in Flux. *Language and Linguistics Compass,* 2/3: 422–441.

Harvey, Mark & Nicholas Reid 1997. *Nominal Classification in Aboriginal Australia.* Philadelphia: John Benjamins.

Harvey, Mark, Nordlinger, Rachel, Green, Ian 2006. From Prefixes to Suffixes: Typological Change in Northern Australia. *Diachronica* 23 (2): 289–311

Ikoro, S.M. 1996. *The Kana language*. Leiden: CNWS.

Jones, Robert B. 1970. Classifier constructions in Southeast Asia. *Journal of the American Oriental Society*, 90(1):1-12.

Matisoff, J. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman, (ed.), *Consonant Types and Tone,* 71-95. *Southern California Occasional Papers in Linguistics,* No. 1. Los Angeles: UCLA.

Mithun, Marianne 1986. The convergence of noun classification systems. In Colette Craig (ed.), *Noun classes and categorization*, 379-397. Philadelphia: John Benjamins.

Paterson, Rebecca 2012. The semantics of Ut-Ma'in noun classes. In R.M. Blench & S. McGill *eds. Advances in minority language research in Nigeria*. *Volume I*. 239-272. Köln: Rudiger Köppe.

Post, M.W. 2007. *A grammar of Galo*. Ph.D. RCLT, La Trobe University, Bundoora.

Reh, M. 1985. *Die Krongo-Sprache (nìino mó-dì)* : *Beschreibung, Texte, Wörterverzeichnis.* Kölner Beiträge zur Afrikanistik, 12. Berlin: Dietrich Reimer.

Reh, M. 1994. A grammatical sketch of Deiga. *Afrika und Übersee*, 77:197-261.

Sagart, Laurent 2011. The Austroasiatics: East to West or West to East? In: N. Enfield ed. *Dynamics of Human Diversity in Mainland SE Asia*. 345-359. Canberra: Pacific Linguistics.

Schadeberg, Thilo C. 1981a. *A Survey of Kordofanian. Volume 1: The Heiban Group*. Hamburg: Helmut Buske.

Schadeberg, Thilo C. 1981b. *A Survey of Kordofanian. Volume 2: The Talodi Group*. Hamburg: Helmut Buske.

Shafer, Robert 1952. Etudes sur l'austroasien. *Bulletin de la Société de Linguistique de Paris,* 48: 111-158.

Shorto, Harry L. 2006. *A Mon-Khmer comparative dictionary*. Paul Sidwell, Doug Cooper and Christian Bauer eds. PL 579. Canberra: ANU.

Sidwell, Paul & Roger M. Blench 2011. The Austroasiatic *Urheimat* : the Southeastern Riverine Hypothesis. In: N. Enfield ed. *Dynamics of Human Diversity in Mainland SE Asia*. 317-345. Canberra: Pacific Linguistics.

Smith, Kenneth D. 1975. The velar animal prefix in Vietnam languages. *Linguistics of the Tibeto-Burman Area*, 2:1-18.

Stevenson, Roland C. 1991. Relationship of Kadugli-Krongo to Nilo-Saharan: Morphology and lexis. In: *Proceedings of the Fourth Nilo-Saharan Conference. Bayreuth, 1989*. Bender, M L. (ed.) 347-369. Hamburg: Buske.

Storch, Anne 2005. *The Noun Morphology of Western Nilotic*. Köln: Köppe.

Thomas, Dorothy. 1969. Chrau affixes. *Mon-Khmer Studies*, 3:90-107.

Van Driem, George L. 1998. Neolithic correlates of ancient Tibeto-Burman migrations. In: *Archaeology and Language II*, R.M. Blench and M. Spriggs eds. 67-102. London: Routledge.

Van Driem, George L. 2008a. To which language family does Chinese belong, or what's in a name, pp. 219-253 in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin, eds., *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*. London and New York: Routledge.