

**Austroasiatic and Sino-Tibetan:
convergence or contact? A
speculative history of 'noun class'
affixes**

**Max Planck Institute for Evolutionary
Anthropology**

Leipzig, November 29th -1st December 2012

Roger Blench

Kay Williamson Educational Foundation

Word structure in SE Asian language phyla

- The Sino-Tibetan and Austroasiatic language phyla, despite being geographically intertwined in SE Asia, are not usually thought of as being genetically related.
- Despite this, they have a strikingly similar word structure, usually known in the regional literature as 'sesquisyllabic'.
- However, words typically have a C prefix and a stem that looks as if it is underlyingly CVCV, though it is shortened in many languages
- In the regional literature the terms 'minor' and 'major' syllable are used
- Chinese, of course is not like this, but proto-Sinitic might well have been; Tibetan does have this structure
- It is tempting to pronounce words as if they had a consonant cluster at the beginning but it ain't so.

Word structure in Sino-Tibetan and Austroasiatic

- Minor syllable prefixes are either lost or optional in many languages, and they seem to be able to change in ways that do not suggest phonological shift but affix substitution.
- In Austronesian languages, the tendency is for the prefix to have an obligatory vowel 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.
- The Daic languages, which are typically CV(C) and except in rare cases no longer retain the minor syllable.

Word structure in Sino-Tibetan and Austroasiatic

- These prefixes do *not* mark number and thus do not alternate but they do have semantic associations; there is a particularly widespread k- prefix in Austro-Asiatic marking animals.
- And an S- prefix associated with animals, occurring in both Austroasiatic and Sino-Tibetan
- Moreover, the prefixes can be exchanged in cross-linguistic perspective, in other words, the stem will remain the same and a new prefix acquired
- So it is reasonable to assume that there was once a much more widespread system of semantically assigned prefixes and that this has eroded, but is still partially present in the minds of speakers
- If so, how did this system originate?

Classifiers in Sino-Tibetan and Austroasiatic

- Sino-Tibetan and Austroasiatic, along with many other language phyla (Austronesian and many New World) are marked by nominal classifier systems
- These are essentially grammaticalised nouns that have become obligatory accompaniments (clitics?) when marking plurals or groups of nouns
- They do not usually show any type of agreement (?exceptions in the Amazon)
- It is likely that the prefix systems in Sino-Tibetan and Austroasiatic are in fact frozen classifiers, which preceded noun stems and then became partially incorporated
- A new classifier would then be applied to the stem
- This very much corresponds to *affix renewal* in Niger-Congo where noun-class affixes become unproductive and a new affix is added

Some examples 'Buffalo' #*krəpəw* in Austroasiatic

Branch	Language	Attestation
	Proto-Mon-Khmer	* <i>krpiʔ</i>
Monic	Mon	preaŋ 𑜀𑜢𑜤𑜰𑜫 𑜀
Monic	Nyah Kur	chəlow
Vietic	proto Vietic	* <i>c-lu</i>
Vietic	Thavung	khuay
Vietic	Pong	klow
Khmeric	Khmer	krəbɿy
Pearic	Pear	krəpa:w
Pearic	Chong	kapa:wA
Bahnaric	PNB	* <i>kapɔ:</i>
Bahnaric	PSB	* <i>g~rəpu:</i>
Bahnaric	Sedang	kopôu
Bahnaric	Tampuon	kəpəu
Bahnaric	Bahnar	kəpo ː
Katuic	Proto-Katuic	* <i>krpiiw</i>

'Buffalo' #*krəpaaw* in Austroasiatic (cont.)

Branch	Language	Attestation
Katuic	Pacoh	kərbɤ:
Katuic	Chatong	karpiw
Khmuic	Sre	rəpu
Khmuic	Chrau	gəpu ː
Khmuic	Biat	rpu ː
Aslian	Kensiw	kɛ'paw
Aslian	Temiar	kəɽbau

'Buffalo' #*krəpaaw* in Austroasiatic (cont.)

- ❖ The assumption is that the original stem was *-pəw* or similar. The primary prefix was *r-* as still survives in Khmuic
- ❖ However, the relation between *k-* prefixes and animals subsequently spread rather irregularly across the phylum (not attested in Munda and Khasian I think)
- ❖ In some cases it replaces the *r-* prefix, elsewhere it becomes a pre-prefix.
- ❖ In Mon, the *r-* prefix moves rightward into the stem (very common in Niger-Congo languages)
- ❖ In Vietic, assuming these forms are cognate, *k* → *tʃ* and *r* → *l* and the bilabial is lost entirely

S- prefix Animal names in Nafra Miji

Gloss	ʃ	s/ts
animal		stɕõ
horse	ʃgrɔ	
sheep	ʃgθɔʔ	
goat	ʃprn	
dog	ʃazi	
deer		tstə
flying squirrel		ʃbiã
leopard	ʃnmũ	
monkey	ʃbõ	
pangolin, anteater		ʃgdʒɔ
wild cat	ʃgrɛ̃	
wild dog	ʃkfə	
sparrow		slĩʔ
ant	ʃni	
fish sp. I		sθw̃

Affix movement and loss in the evolution of Chinese

Gloss	Early Zhou	Classical	Character
blood	s.wi:t	wi:t	血
fire	s.mə:y?	mǎ:y	火
head	s.lu?	lǔ	首
black	s.mə:k	mǒ:k	黑
see	ke:n.s	ke:n.s	見
seed	toŋ.?	tǒŋ	種

Contact, borrowing and metatypy I

- ❖ 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.

Contact, borrowing and metatypy II

- ❖ 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.

Contact, borrowing and metatypy III

- ❖ Both Austroasiatic and Sino-Tibetan may originally have had simple stems, with no affixes marking number, case, semantics or gender. Nominal classifiers, usually CV(C) syllables with semantic assignments, and were put together with nouns, usually preceding them.
- ❖ In Sino-Tibetan and Austroasiatic nominal 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, elsewhere a new prefix was added in front of the existing prefix. Meanwhile, distinct nominal classifiers continued to co-exist and continued to be incorporated and renewed.

This can be taken to demonstrate

- ❖ Sino-Tibetan and Austroasiatic have underlyingly similar word structures, without being genetically related
- ❖ 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
- ❖ that unproductive affixes can be subject to renewal, for example reprefixing, without forming consonant clusters
- ❖ that semantic associations of affixes both can be borrowed across phyletic boundaries, along with the segmental material, and evidently were at an early stage of the evolution of these phyla

Noun classes: Niger-Congo and elsewhere

- Globally speaking, Niger-Congo type noun-classes are fairly rare; semantically based systems of number marking with affixes and concord on adjectives and other parts of speech really only occur in Africa in parts of Niger-Congo
- Indeed, they are not even typical of the whole of Niger-Congo, despite a large and careless literature to the contrary.
- They are lacking in Mande, Dogon, Ijoid, Kaalak-Domurik & ?Rashad apparently from the beginning
- They appear to have eroded in much of Kwa and Volta-Niger
- So it is quite likely that they appeared partway through the evolution of Niger-Congo and the non-class languages are at the top of the tree..

Noun classes: Niger-Congo and elsewhere

- Elsewhere in the world, the main area where these occur is in Papuan and Australian languages. In most cases these languages have only three or four classes, but a few have evolved complex systems comparable to Niger-Congo
- Similar systems without the same type of concord appear in North Caucasian and Yeniseian
- Nilo-Saharan has striking systems of affix alternation marking number in a few branches, Daju, Kadu and Koman, but these are not associated with semantics or concord
- These can often be associated with a three-term system of number-marking
- So it seems a reasonable question to ask how noun-classes evolved.

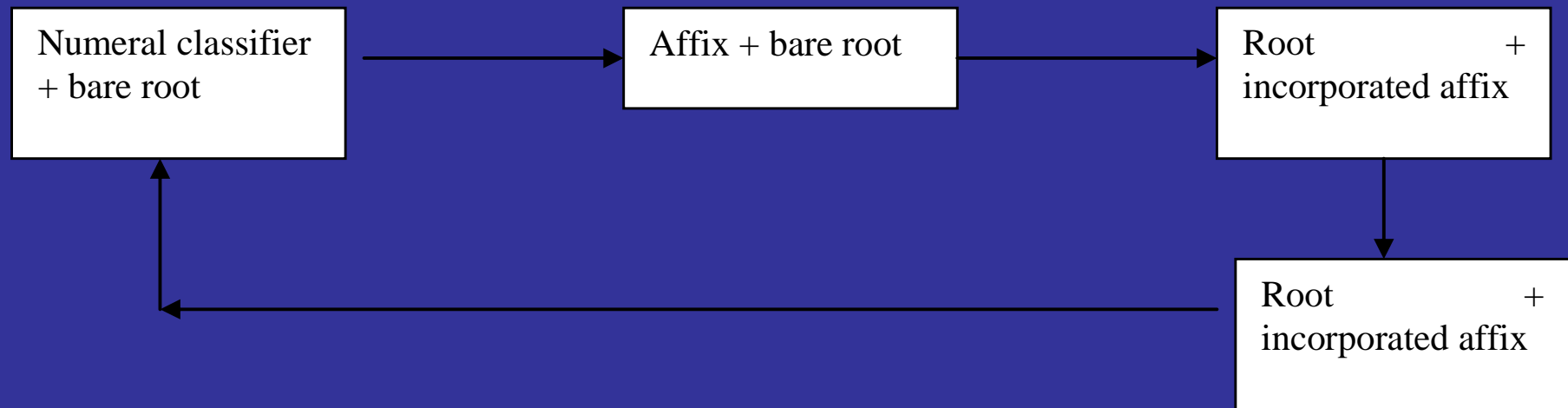
Classifiers in Africa? I

- Nilo-Saharan languages does not have concord but does have productive affixes and affix renewal
- A language like Krongo can have up to three frozen affixes
- Which of course is part of the reason Greenberg classified the 'Tumtum' languages as Niger-Congo
- We do not usually consider African languages as having nominal classifiers, or SE Asian languages as having noun-classes.
- But there is increasing evidence for the secondary evolution of nominal classifiers in Niger-Congo languages
- The most well-known case is Kana, an Ogoni language, part of the Cross River group.
- Kana has pretty much lost its nominal morphology, and the classifiers, may be an attempt to compensate, as it were

Classifiers in Niger-Congo? II

- Other examples have been mooted, such as in Ejagham (Ekoid), although rarely described in detail
- An interesting example is Mambay, an Adamawa language spoken in north-central Cameroun, described by Erik Anonby.
- Mambay has a functioning system of noun-class suffixes, but which appears to be developing prefixed classifiers
- Examples given by Anonby include the 'collectives' which precede nouns, do not show concord and have broad semantic correlations
- I suspect these systems are more common than has been recognised, as a function of what we expect to find in various language phyla.

Cyclical renewal of affixes in SE Asian languages



A historical scenario I

❖ 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. 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.

A historical scenario II

❖ 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 river valleys to grow taro while using improved boat technology. This would have been around four thousand years ago, when there is a rapid and sudden expansion of the Neolithic in mainland SE Asia

❖ 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.

Conclusion

- ❖ 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 piled 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.

THANKS

- To Kay Williamson Educational Foundation for supporting the fieldwork
- And many people who helped in Arunachal Pradesh, Nagaland, Laos and Nigeria

