References

Ballard, W.L. 1969. Phonological History of Wu. Ph.D. thesis, University of California, Berkeley.

Brown, Marvin. 1985. *From Ancient Thai to Modern Dialects*. Bangkok: White Lotus Co.

CH. 1964. 汉语方言词汇 *Hanyu Fangyan Cihui* [Dictionary of Vocabulary in Chinese Dialects]. Peking: Wenzi Gaige Chubanshe.

Cao Z 曹志耘. 2005. 南部吴语语音研究 [Studies in the Phonetics of the Southern Wu dialects]. Peking: Shangwu Press.

Chao Yuen Ren. 1928. 現代吳語的研究 / *Studies in the Modern Wu Dialects*, Monograph 4. Peking: Tsing Hua College Research Institute.

Fu Guotong傅国通, Fang Songxi, 方松熹 Cai Yongfei 蔡勇飞, Bao Shijie 鲍士杰 & Fu Zuozhi 傅佐之.1985. 浙江吴语分区 [Sub-groups of the Wu dialects of Zhejiang]. Zhoushan: Zhejiang Linguistics Society.

Maddieson, Ian. 1978. "Universals of Tone". *Universals of Human Language* ed. by J.H. Greenberg, vol. 2: Phonology, 335-66. Stanford: Stanford University Press.

Norman, J. 1988. *Chinese*. Cambridge: Cambridge University Press.

Rose, Phil. 1987. "Considerations in the normalisation of the fundamental frequency of linguistic tone". *Speech Communication* 6:4.343-352.

Rose, Phil. 1990. "Acoustics and Phonology of Complex Tone Sandhi". *Phonetica* 47:1-35.

Rose, Phil. 1993. "A Linguistic Phonetic Acoustic Analysis of Shanghai Tones". *Australian Journal of Linguistics* 13.185-219.

Rose, Phil. 2000. "Hong Kong Cantonese Citation Tone Acoustics: A Linguistic-Tonetic Study". *Proceedings of the 8th Australian International Speech Science and Technology Conference* ed. by Michael Barlow, 198-203 Canberra: Australian Speech Science and Technology Association.

Rose, Phil. 2001. "Wu", *Facts about the World's Languages* ed. by J. Garry & C. Rubino, 158-161. New York: New England Publishing Associates.

Rose, Phil. 2002. "Independent depressor and register effects in Wu dialect tonology: Evidence from Wenzhou tone sandhi". *Journal of Chinese Linguistics* 30:1.39-81.

Rose, Phil. 2004. "Zooming-in on Oujiang Wu: tonal homogeneity and acoustic reconstruction in a small sub-group of Chinese dialects". *Proceedings of the 11th Australasian International Conference on Speech Science and Technology* ed. by C. Watson & P. Warren, 58-63. Auckland: Australasian Speech Science and Technology Association.

Wurm, S.A. T'sou, B. Bradley, D. Li R, Xiong Z. & Zhang Z., eds. 1987. *Language Atlas of China*. Canberra: Longmans/ANU Research School of Pacific Studies.

ZH. 1989 (2nd edition). 汉语方音字汇 *Hanyu Fangyin Zihui* [A dictionary of pronunciations of characters in Chinese dialects]. Peking: Wenzi Gaige Chubanshe.

Zhengzhang S. 1987. "Wu Group". Wurm et al. 1987. (no page nos.)

Zhu S. Xiaonong. 1999. *Shanghai Tonetics*. Munich: Lincom.

Zhu S. & Rose, Phil. 1998. "Tonal complexity as a dialectal feature: 25 different citation tones from four Zhejiang Wu dialects". *Proceedings of the 5th ICSLP* ed. by R. Mannell & J. Robert-Ribes Vol 3, 919-922. Sydney: ASSTA.

ISSUES IN THE MORPHOLOGICAL RECONSTRUCTION OF PROTO-MON-KHMER

PAUL SIDWELL
Australian National University, and
Centre for Research in Computational Linguistics (Bangkok)

1. *Introduction*[1]

The honouree of this volume is famed for his passionate interest in morphology, morphological reconstruction and his championing of etymological methods in reconstruction. In this context, it gives me great satisfaction to honour Harold by discussing some issues in the morphological reconstruction of Proto-Mon-Khmer, and setting forth a reconstruction of various Proto-Mon-Khmer affixes.

The Mon-Khmer family belongs within the Austroasiatic phylum of South and Southeast Asia. The morphological systems of these have been described to greater and lesser extents, and comparative Mon-Khmer studies have been conducted - with stops and starts - for more than century already. It is now clear that Proto-Mon-Khmer bound morphology was characterised (at least) by prefixing and infixing, employed for purposes connected with derivation and verbal semantics. There was also much creativity in the formation of expressive lexicon, with much use of reduplication and assimilation. The latter are not discussed here; instead, it is only affixes that will be examined.

Approaches to Mon-Khmer morphological reconstruction have faced various hurdles; one being that since the foundational studies of Schmidt (especially 1905, 1906) the issues have tended to be seen through the lens of the Austric hypothesis. As formulated by Schmidt, Mon-Khmer and Austronesian are sisters within an Austric macro-phylum, and although many lexical parallels have been adduced, the most important evidence has always been seen to reside in the morphology, specifically in among a small subset prefixes and infixes (see Reid 2004, 1999, 2005 for recent views). Thus, much of the modest relevant research effort has effectively presumed that the Mon-Khmer-Austronesian morphological parallels are archaic and reflect reconstructable morphemes. Such a top-down approach to reconstruction has its own merits, but in the present case we are dealing with very old languages families whose affilitation is not generally accepted, and may well be impossible to demonstrate. In this

paper I advocate a more bottom-up approach, emphasising the comparative analysis of forms which occur within a well-established genetic grouping, and which for formal and distributional reasons are likely to be archaic, and not borrowed.[2]

2. *Methodological preliminaries*

Mon-Khmer languages have a distinctive morphological typology, characterised by prefixation and infixation of mono- and sesquisyllabic stems. The preferred phonological shape of unaffixed Proto-Mon-Khmer stems (based on Shorto 2006) included:

- Monosyllabic: C(R)V(V)C, where R was an optional medial consonant, one of the set /r, l, j, w, h/, and:
- Sesquisyllabic: CCV(V)C, in which initial clusters of stops (oral and nasal) were permitted, with a weak vowel epenthesis after the first C, creating a light initial syllable or 'presyllable'.

Initial sequences of two or more consonants could be built up by affixation, with a likely preference for initial sequences of three consonants in morphologically complex words, affording presyllables a coda if the stem was light (CVC) (Anderson (2004) has a similar formulation he calls the "bimoraic constraint"). Thus, some affixes had mono- and bi-consonantal allomorphs, applied according to the weight and/or sonority contour of the stem.

Syntactically the languages are isolating, so that morphological processes are largely restricted to derivational functions, although one finds quite a variety of idiosyncratic developments – for example the case marking of pronouns of (Katuic) Ta'oih (Solntseva 1996) and Pacoh (Alves 2006) are transparently innovative.

Not withstanding the above, speakers of many Mon-Khmer languages have more or less abandoned affixation as a morphological strategy, typically in the context of phonological changes that have restructured the lexicon into monosyllables. In these languages, such as Vietnamese, Nyaheun, and others, compounding has emerged as an important new word-formational mechanism (probably for both internal structural reasons, plus contact with compounding languages such as Chinese, Thai, Lao).

---

[2] One of the reviewers of this paper suggested that I should consider the possibility that various items of Mon-Khmer bound morphology may be ancient borrowings from Austronesian. My intuitive response is to dismiss the suggestion as unlikely and difficult to test, although I must admit that it is not impossible. Probably such a hypothesis cannot be tested properly without first conducting far more extensive bottom-up studies of Austroasiatic morphology.

Some reconstructions of affixal forms have already been suggested. Relying especially upon Nicobarese evidence, Reid (1994) offers the following Proto-Austroasiatic affixes:

Causatives:	*pa/-ap-, *ka-
Agentives:	*-um-, *ma/-am-
Instrumental:	*-an-, *-in-
Objective:	*-a

The last of these, the Objective suffix, is especially doubtful. The form does not occur elsewhere in Austroasiatic; the Nicobarese "suffixes" are analysed as clitics by Braine (1970), and should be considered innovations. The other affixes reconstructed by Reid do have cognates in other branches, and therefore do represent ancient morphemes. However, placing weight on Nicobarese led Reid to reconstruct these affixes with CV/VC structure, even specifying the quality of the vowels. This may be correct for Nicobarese, but it is quite possible that they were purely consonantal in Proto-Mon-Khmer, and may have been similarly so in Proto-Austroasiatic. Compare with Anderson (2004) who, giving special weight to Munda forms, reconstructs the Proto-Austroasiatic causative as a prefix *'b- with CVC stems and as an infix *-'b- with CCVC stems.[3] The uncertain place of Nicobarese in the Austroasiatic family, and the lack of a well-developed comparative reconstruction of Proto-Nicobar, means that the use of such evidence for wider comparative analysis is extremely problematic. Clearly there are issues of segmental and prosodic phonology that are far from solved at the Proto-Austroasiatic level, and for now I prefer to focus on reconstruction at a lower phylogenetic level.

Taking the bottom-up approach with the Mon-Khmer family, it is convenient to begin by identifying at least two branches of the family that maintain evidence of rich morphological systems, and yet are widely separated geographically. Guided by these principles we would turn immediately to the Aslian languages which, located on the Malay Peninsula, have escaped the areal effects of contact with Thai/Lao that affected so much of the mainland (pointed out by Diffloth 1979). Fortuitously, the Aslian languages show a morphological complexity which has been well described, and in this study, I draw upon Jahai (Burenhult 2005) and Semelai (Kruspe 2004).

At the geographical extreme from Aslian are the Northern Mon-Khmer languages: Khasi, Palaungic and Khmuic. The first two are not promising; Palungic languages have little morphology to offer, and "morphological processes in Khasi are almost non-existent" (Rabel-Heymann 1976:981). However, Khmuic is a promising candidate for morphological comparison, since documented varieties show considerable morphological complexity (although only modest morphological productivity). Thanks to apparently very

---

[3] I assume this notation represents *ɓ-/-ɓ-, since Anderson describes the sounds as "glottalised".

conservative phonology, extensive complex forms persist, which aid in the historical analysis. For this study I draw upon the Kammu documented by Svantesson (1983) and the Khmu of Suwilai (2002).[4]

Svantesson anticipates the selection of Aslian and Khmuic as criterion branches for morphological reconstruction:

> It is believed that Proto-Mon-Khmer (and Proto-Austroasiatic) had a rich system of derivational affixes (Pinnow 1960, Diffloth 1976), and some of the affixes found in Kammu (e.g. causative /p-/ are widespread in different Austroasitic subgroups. Aside Munda, the Austroasiatic subgroups having the most developed morphological systems are probably Aslian (see e.g. Benjamin 1976a) and Nicobarese. In Vietnamese there is almost no derivational morphology at all, and in most other groups, the situation is similar to that in Kammu, i.e. affixation is common but not completely productive. (Svantesson 1983:73)

There are also other factors, in addition to morphological complexity, that point to both Khmu and Aslian being generally rather conservative as Mon-Khmer languages go. Many Mon-Khmer languages that have been in contact with Tai have undergone a distinctive type of phonological restructuring, in which voiced stops devoiced, and the vocalism was restructured to double the number of contrasts, either by acquiring tones, breathy phonation, a plethora of diphthongs, or a combination of two or three of these factors. Typically the Proto-Mon-Khmer imploded stops continue as a distinct series, often losing implosion to become the 'new' voiced series (as happened in Thai and Lao). This did not occur in Aslian languages, which maintain the old voicing contrast (in fact merging the Proto-Mon-Khmer imploded stops into the voiced series), instead many of them lost the length contrast in the vowels, apparently under the influence of Austronesian. Many of the Khumic languages, although spoken in the zone of prolonged Tai contact, have retained the Proto-Mon-Khmer voiced series (also merging the Proto-Mon-Khmer imploded stops into the voiced series as did Aslian), and have resisted vowel restructuring. Suwilai Premsrirat's Khmu Chuang is such a conservative variety.

Svantesson (1983:83) lists some 27 distinct forms of "affixators" in Kammu. Several of these are transparently innovative forms or specifically related to expressive formations, and these are removed from consideration for now. The remaining forms are reorganised according to function and listed as follows (prefixes indicated by x-, infixes indicated by -x-, allomorphs by /):

Verb forming:	gn-, kn-, km-
Resultative:	tl-, sl-, hn-
Causative:	p-, pn-, -m-
Reciprocal:	tr-
Nominalising:	rn-, -rn-/-r-/n(d)-, -n-, -mn-, sŋ-, sr-, tr-

---

[4] Henceforth forms noted as 'Kammu' are all from Svantesson (1983), and those noted as 'Khmu' are from Premsrirat (2002).

Prefixes are straightforwardly added to the left edge of stem, while infixes are inserted to the right of the leftmost consonant. Speakers do not readily tolerate initial sequences of more than three consonants, so there are consequent limits on the amount of affixation that may applied, and in some cases there is deletion of segments to accommodate affixation.

A remarkable phonological concomitant of Kammu affixation is what Svantesson calls "coda assimilation". Affixation tends to creates presyllables, and in certain cases, the rhymes of these presyllables (often underlyingly /n/ or /r/) assimilate to the coda of the mainsyllable, e.g. (from Svantesson 1983:58):

Nominalizing rn-:

hɔ́ɔc 'to carve'	>	rc̓hɔ́ɔc 'carving'
wàaj 'to swim'	>	rĭwàaj 'fin'

Causative pn-:

kák 'bent'	>	p̌kkák 'to bend'
rɔ̀os 'angry'	>	pšrɔ̀os 'to make somebody angry'

The situation in Aslian is more complex; in particular, there is a far more elaborated version of coda assimilation, which Aslian specialists variously call "coda copying", "reduplicative infixation", "incopyfixation"[5] The difference is that instead of a well-formed segment assimilating to the mainsyllable coda, it is an underspecified or epenthetic segment (or segments) that assimilates. Consider the imperfective aspect marker in Jehai (from Burenhult 2005:94-5): for monosyllabic stems the prefix consists of a copy of the mainsyllable onset plus a copy of the mainsyllable coda, e.g.:

sut 'to sob'	>	stsut 'sobbing'
weʔ 'to exist'	>	wʔweʔ 'existing'

For sequisyllabic stems the allomorph is simply a copy of the mainsyllable coda inserted to create a minorsyllable rhyme, e.g.

tʔoc 'to ask'	>	tcʔoc 'asking'
rwis 'to cut grass'	>	rswis 'cutting grass'

In disyllabic stems[6] the minorsyllable rhyme gains a coda:

turɛh 'to tap poison'	>	tuhrɛh 'tapping poison'
gulɛm 'to carry'	>	gulɛm 'carrying'

---

[5] This last term coined by coined by Matisoff (2003:28) with charateristic wit.

[6] Disyllabic stems are not a general feature of Mon-Khmer languages, and in Aslian can be assumed to have arisen under Austronesian influence. Many in fact are obvious Malay loans.

Note that the allomorphy maintains the (historically?) preferred word-shape of morphologically complex forms. This allomorphy also means that the full set of Jehai affixes is difficult to summarise abstractly. Here is table of verbal derivation from Burenhult (p.94):

Derivational morpheme	Monosyllabic /cip/	Sesquisyllabic /kjeŋ/ 'to listen'	Disyllabic /gulɛm/ 'to carry'
Aspect/Aktionsart			
Imperfective	cp-cip	k<ŋ>jeŋ	gu<m>lɛm
Progressive	b-cpcip	b-kŋjeŋ	b-gulɛm
Iterative	lp-cip	l-kjeŋ	l-gulɛm
Continuative	cip-cip	kjeŋ-kjeŋ	gulɛm-gulɛm
Distributive	cip-cip	k<iŋ>jeŋ	g<im>lɛm
Reciprocal	ca-cip	k<a>jeŋ	—
Affix /m/	mp-cip	mŋ-jeŋ (?)	mm-lɛm (?)
Causative	pp-cip, pi-cip, pr-cip, tr-cip	k<ri>jeŋ	g<ri>lɛm
Nominalisation	np-cip	k<nŋ>jeŋ	g<n>ulɛm
Collective plural nominalisation	—	k<ra>jeŋ	g<re>lɛm (?)
Relative	t-cip (?)	t-kjeŋ (?)	t-gulɛm (?)

Figure 2: *Paradigm of verbal derivation in Jahai (Burenhult 2005:94, Table 4.5)*

So what is the historical context of Aslian incopyfixation? Matisoff suggests that it is not especially unique:

> Such processes are still apparently productive in Nancowry Nicobarese (Radhakrishnan 1970:149ff). A few Proto-Waic reconstructions look like they have fossilized incopyfixes, e.g. *rmhom 'heart, mind' [Diffloth 1980, p. 148]. Delcros' Khmu dictionary (1966) has forms like *rtjut* 'horrible' *ɭnbaaɲ* 'immense'. (fn.150, p.30)

My theory is that the origin is phonological, and was connected to the structure of sesquisyllables and associated phontactic constraints. I expect that the Kammu pattern is more-or-less indicative of the Proto-Mon-Khmer state of affairs; coda assimilation was not necessarily especially productive, and many of the forms still in use are fossils, the accumulated detritus of centuries of sporadic formations. On the other hand, the Aslian languages, once separated from other Mon-Khmer speakers, went on to regularise and elaborate, with innovations such as the phonologically underspecified morphemes. At the earliest stage (at the Proto-Mon-Khmer level?[7]) speakers inserted epenthetic rhymes into consonant clusters, and with euphonic/expressive motivations began to rhyme these with the mainsyllable codas. This is quite consistent with the sorts of word game that are common in Southeast Asia (e.g. swapping

---

[7] I haven't (yet) found Aslian-type epenthetic coda copying described in respect of Munda languages, although expressive reduplication is common to them.

rhymes within couplets of monosyllables is popular in Vietnamese), and expressive reduplication is described in detail for many Mon-Khmer languages.

In the following section, I will present evidence for the reconstruction of a selection of Proto-Mon-Khmer affixes. The presentation is far from comprehensive, dues to space considerations, and should be taken as illustrative rather than authoritative.

3. Infixes

3.1 Mon-Khmer Infixes

Infixation is an important word-formational process in Mon-Khmer languages. It has been suggested that generally infixes can be analysed as a type of prefixation, a position that Kruspe takes in her analysis of Semelai, arguing that the infixes are effectively prefixed to the prosodic head (2004:65-6). Looking for an historical explanation one may observe that among Mon-Khmer languages the most common infixal segments are nasals, while nasal onsets in sesquisyllabic stems are rare (or nonexistent). By contrast, nasal codas are common in all types of mainsyllables - the proportion typically approaches 50% (except in those languages in which hardening of finals is a regular development). A similar but weaker tendency applies to liquids. This suggests that during a specific historical stage of Mon-Khmer (pre-Proto-Mon-Khmer ?) nasals and liquids were not permitted as leading onset segments of sesquisyllables. Whatever the particulars of this constraint, it may have been associated with a metathesis of sonorant initials, i.e. NC- > CN-. Perhaps we see an echo of just such a process in the behaviour of Jehai affixes (see examples under 3.3).

Another point to consider is whether there is an historical relation between /n, m, r/ as the typically infixal segments, and the implosives /ɓ, ɗ/. There is a remarkable symmetry in the fact that the implosives are only reconstructable for Proto-Mon-Khmer as prevocalic segments, but never preconsonantally (in contrast, say, to contemporary Bahnar, see Banker et al. 1979). A general lenition of preconsonantal implosives, followed by metathesis (**ƊCVC > **NCVC > *CNVC) would account for the existence and phonological form of Austroasiatic infixes. It this context it is intriguing that Anderson's (2004) reconstruction of the Munda causative suggests an implosive articulation.

3.2 Nominalising *-n-

The nominalising *-n- is distributed widely, although in some languages it is no longer productive, or is restricted in its productivity (e.g. only weakly productive in Khmuic). It appears that essentially any verb could be nominalised with *-n-. It is not clear if its usage was originally restricted (to perhaps

instrumentals or resultatives) and subsequently generalised, or the other way around.

Kammu (instrumental):

kóh	'to cut'	knóh	'cutting board'
kèɛp	'to fasten belt'	knèɛp	'quiver'

Khmu (resultative):

pɛɛr	'to slice'	pnɛɛr	'sliced pieces'

There are vestiges in other Northern Mon-Khmer languages, such as Khasi (data from Henderson 1976):

Khasi:

shong	'to sit, dwell'	shnong	'place, village, town'
sait	'to wash (vegetables)'	snait	'strainer'

In Aslian the affix is accompanied by a copy of the mainsyllable coda:

Jahai:

CVC stem	sam	'to hunt'	nmsam	'act of hunting'
CCVC stem	tbɔh	'to beat'	tnhbɔh	'act of beating'

Semelai:

CVC stem	cʰɔr	'treat with fire'	nrcʰɔr	'firing'
CCVC stem	smaɲ	'to request'	snɲmaɲ	'request for sth'

The infix shows considerable allomorphy in various other Mon-Khmer languages, essentially relating to restrictions on the occurrence and collocations of /n/. E.g.:

Katu:

	cha	'to eat'	chana	'food'
n > a / _l	pleh	'to turn on road'	paleh	'crossroads'
n > a / _r	pruung	'to blow fire'	paruung	'pipe to blow fire'

3.3   *Nominalising Agentive *-m-*

Aslian provides evidence of a labial nasal agentive affix, e.g.:

Jahai:

wĩ	'left'	mwĩ	'left-handed person'
tem	'right'	mmtem	'right-handed person'

Semelai:

kbɔc	'fish with pole'	kmbɔc	'fishing pole'
sdər	'to remember'	smrdər	'a good rememberer'

In Old Khmer we find a rare 'agentive' -m-, e.g.:

Old Khmer:

siŋ	'to reside, preside'	smiŋ	'in residence; to officiate'

A similar morpheme is found in Mon, and Diffloth (1984) reconstructs a Proto-Monic *-m- agentive. However, the closest match in Kammu/Khmu is an unproductive complex infix -mn- with an instrumental function:

Kammu:

tìs	'to support'	tmínìs	'support'
kées	'to fence in'	kmínées	'fence'

Khmu:

koh	'to chop'	kmnoh	'chopping block'

Perhaps the Khumic -mn- is a remnant of a combination of two infixes, the labial reduced to a remnant of the Proto-Mon-Khmer agentive.

3.4   *Expressive of repetitiveness/numerousness *-l-*

Matisoff (2003:27) notes that:

> Jah Hut has an -l- infix that occurs in expressives to indicate intensity or great numbers (e.g. saʔbyʉr 'sight of dishevelled hair', slaʔbyʉr 'long, abundant dishevelled hair'), and which also occurs fossilised in the form -lʔ- in names of animals characterised by rapid, jerky movements (klʔbak 'butterfly', hlʔdɛʔ 'cockroach', klʔjɛh 'kind of small bird').

One finds similar forms in Khmu:

Khmu:

klbaa	'wave'
kliəʔ	'fluctuating movement'
klwaʔ	'echo, reverberating sound'
klʔus klʔas	'cluttered'
tlgɔɔk	'clf. for cluster of fruit'
tlsuəl	'cluttered (small things)'

And in other Mon-Khmer languages, e.g. compare Sre (South Bahnaric) klʔeep 'centipede' with Khmer kʔaep 'centipede'. This affix may also be the source of the Khmer iterative infix -L-, which derives both nouns and verbs (discussed by Jenner & Pou 1980-81 p.xlix), e.g.:

Khmer:

| toəm | 'to bear' | troəm | 'to support patiently' |
| kǐək | 'to hold against side or under the arm' | klǐək | 'armpit' |

3.5 Nominalising instrumental *-p-

This particular infix is very archaic; it is no longer productive in most Mon-Khmer languages, so discussions are absent from descriptive grammars, and reconstruction is dependent on comparative and philological evidence (hence the reliance on Mon and Khmer). Bauer (1988), comparing Old Mon and Old Khmer reconstructs Proto-Mon-Khmer -p- 'nominaliser'.

Reflexes of *-p- are well attested in Khmer, e.g.:

Khmer:

| dɑl | 'to pound' | tɓɑl | 'mortar bowl' |
| rɛɛŋ | 'to plait' | rɓaeŋ | 'basketry' |

Cognates of the infixed "mortar" etymon are widespread in Mon-Khmer, e.g. (Shorto (2006) entry 1757): Praok (Wa) pɔ, Rumai (Palaung) mpae, Katu tapal, Nyaheun dwaw, Sre mpal, Semang pɛl and others showing the effects of regular historical sound changes.

Old Mon (Shorto 1971:xxiv) also possessed the same infix, which by regular sound change took the form -w-, e.g.:

Old Mon:

| til | 'to plant' | twil | 'cultivable land' |
| pa | 'to do' | puwa | 'deed' |

4. Prefixes

4.1 Mon-Khmer prefixes

Comparing prefixes across Mon-Khmer languages is extremely problematic, as the published descriptions typically treat phonologically disparate forms as allomorphs, without further analysis or explanation. In this short study, I will focus on just a few in an attempt to tease out the historically underlying forms.

4.2 Causative *p-

This prefix has attracted the attention of various scholars, especially as it is well attested in Munda, where it shows an infixal allomorph that is paralleld in Mon-Khmer (see Anderson 2004 for a recent discussion). Within Mon-Khmer the most unmarked reflexes have the form of a voiceless bilabial stop, hence the reconstruction as *p-.

In Kammu the affix is typically attached to intransitive and stative verbs. It has several allomorphs with specific phonotactic distributions: p-, pn-, -m-. In the following examples note that examples of pn- are subject to coda assimilation:

Kammu:

p-	háan	'to die'	pháan	'to kill'
p-	kàa	'to climb'	pkàa	'to cause to go up'
pn-	póol	'to hang (intr.)'	pɳpóol	'to hang (tr.)'
pn- + coda assim.	pàaɲ	'drunk'	pɲáaɲ	'to make drunk'
-m-	skár	'straight'	smŕkár	'to straighten'
-m-	r̀kèŋ	'stretched'	rm̀kèŋ	'to stretch'

In Jahai this prefix has an obligatory augment, either a presyllable coda copied from the mainsyllable coda or a vowel /i/ or syllabic consonant /r/. While it is applied to both transitive and intrasitive verbs, it appears from Burenhult's data that intransitive bases predominate, and adding the preffix does not especially increase transitivity, regardless of the base, but may simply indicate "rather specific and restricted meanings" (2005:106).

Jahai:

p- + copy coda	gej	'to eat'	pjgej	'to feed'
pi-	muc	'to eat'	pimuc	'to feed'
pi- + restricted	kap	'to bite'	pikap	'to tear apart with teeth'
pr-:	hit	'tremble'	prhit	'cause smwn. to tremble'

For Semelai, Kruspe describes a causative prefix par- which has:

- various allomorphs due to dissimiation of the coda
- an allomorph -r- where the stem already has a full presyllable
- an infrequent allomorph p- which Kruspe assumes is borrowed, since it is difficult to suggest motivation for its distribution

The prefix is applied to transitive, intransitive and stative verbal bases. Kruspe (her table 5.4) also lists it as occuring with nouns, but this is not exemplified in her text.

Semelai:

par-	tut	'to blow'	partut	'cause smth. to blow'
pan-	ca	'to eat'	panca	'to feed'
pa-	lɛm	'pleasant'	palɛm	'to cheer up smwn.'
-r-	hampɔŋ	'lightweight'	harpɔŋ	'to lighten'
p-	jʔjiʔ	'dirty'	pjʔjiʔ	'to dirty smth.'

For Car Nicobarese, Braine (1970:162) describes a causative prefix *ha-*, with allomorphs *ʔa-* occurring with laryngeal initial CVC stems, and *mi-*, *-um-* with CVCVC stems. Word initial /h/ goes back to *p by regular sound change in Nicobarese, so it is straightforwardly identified with the Proto-Mon-Khmer *p- prefix.

Car:

ha-	rá·c	'hot'	hará·c	'to heat'
mi-	huriŋ	'black'	miriŋ	'to make black'
-m-	siřóŋkú·	'ashamed'	sumřóŋkú·	'to shame'

In Katu the prefix can be affixed to both transitive and intransitive verbs, and to stative verbs or even nouns to create a causative verb with the meaning "activity that results in state/creation of thing". Examples:

Katu:

verb trans.	cha	'to eat'	pacha	'cause to eat'
verb intrans.	mut	'to run'	pamut	'cause to run'
noun	atet	'rust'	patet	'to make to rust'

On the strength of the evidence presented here, one could suggest that the *pr/par-* forms attested in Aslian have been influenced by the *par-* causative prefix of Austronesian (via Malay?), so that reconstruction of the /r/ augment to Proto-Mon-Khmer is not indicated. However, the augmented forms have near parallels in Khmu, consistent with a prosodic preference for presyllables with codas affixed to CVC roots.

4.3   Reciprocal *tr-

In Kammu the prefix *tr-*, with allomorph *tŋ-*, forms reciprocal verbs, e.g.:

Kammu:

| pók | 'to bite' | tr̀pók | 'to bite each other' |
| ràac | 'to scratch' | tr̀ràac | 'to scratch each other' |

Costello (1966:70) documents a Katu *ta-* reciprocal prefix, e.g.:

Katu:

| kap | 'to bite' | takap | 'to bite each other' |
| leng | 'to kill' | taleng | 'to kill each other' |

This prefix, with effectively the same meaning, form and distribution is also attested in Bahnaric languages such as Jeh (Gradin 1976) and Bahnar (Banker 1964).

In various Aslian languages this affix appears to have been replaced by an Austronesian borrowing, apparently Malay *ber-*, e.g.: Semelai *b(r)-* characterised by Kruspe as marking middle voice (fn.6, p117) which includes reciprocity. For Temiar, Benjamin (1976:179-80) describes an equivalent morpheme *bar-* with a similar range of meanings. According to Burenhult reciprocal is marked in Jehai by an *-a-* infixed to the verb.

5.   Closing remarks

Presently the comparative-historical reconstruction of Proto-Mon-Khmer is still grappling with important issues of phonology, lexicon, and classification. Nonetheless, it is evident that much can be achieved in terms if comparative reconstruction of affixes and morphological processes, and the time is right for a comprehensive bottom-up tabulation and comparative analysis of Mon-Khmer morphology.

References

Alves, Mark. 2006. *A Grammar of Pacoh: a Mon-Khmer language of the central highlands of Vietnam*. Canberra: Pacific Linguistics.

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

Banker, John, Elizabeth Banker & Siu Mo'. 1979. *Bahnar Dictionary, Plei Bong-Mang Yang Dialect*. Huntingdon Beach: Summer Institute of Linguistics.

Bauer, Christian. 1988. "Khmer Nasal Affixes: Old Mon Borrowing or Proto-Mon-Khmer Retentions?" *The International Symposium on Language and Linguistics* ed. by Cholticha Bamroongraks, Wilaiwan Khanittanan, Laddawan Permcharoen, Sarinee Anivan, Nancy Bishop, Gwyn Williams, Deeyu Srinarawat, Sirinee Chenvidyakarn, Pornitipa Thongsawang, Mary Peterson, John Hinds, Yuphaphan Hoonchamlong & Nantana Ronakiat, 248-257. Bangkok, Thailand: Thammasat University.

Benjamin, Geoffrey. 1976. "An outline of Temiar grammar". *Austroasiatic Studies* ed. by Philip N. Jenner, Laurence C. Thompson, & Stanley Starosta, 129-188. Honolulu: University of Hawaii Press.

Braine, Jean Critchfield. Nicobarese Grammar (Car Dialect). PhD thesis, University of California, Berkeley.

Burenhult, Niclas. 2005. *A grammar of Jahai*. Canberra: Pacific Linguistics.

Chazée, Laurent. 1999. The Peoples of Laos: Rural and Ethnic Diversities. Bangkok: White Lotus.

Costello, Nancy A. 1966. "Affixes in Katu". *Mon-Khmer Studies* 1.63-86.

Delcros, Henri and Jean Subra. 1966. *Petit dictionnaire du langage des Khmu' de la région de Xieng-Khouang*. Vientiane: Mission Catholique (Mimeograph).

Diffloth, Gérard. 1974. "Austro-Asiatic Languages". *Encyclopaedia Britannica*, 480-484. Chicago: Encyclopaedia Britannica Inc.

Diffloth, Gérard. 1979. "Aslian languages and Southeast Asian prehistory". *Federation Museums Journal* 24.3-16.

Diffloth, Gérard. 1980. "The Wa Languages". *Linguistics of the Tibeto-Burman Area*. 5:2.1-182.

Diffloth, Gérard. 1984. "The Dvaravati-Old Mon Language and Nyah Kur" *Monic Language Studies 1*. Bangkok: Chulalongkorn University Press.

Diffloth, Gérard. 2005. "The contribution of linguistic palaeontology to the homeland of Austro-asiatic". *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics* ed. by Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas, 79-82. London: Routledge/Curzon.

Dournes, Jacques. 1950. *Dictionnaire Srê (Köhö)-Français*. Saigon: des Missions Etrangères de Paris.

Henderson, Eugénie J.A. 1976. "Vestiges of Morphology in Modern Standard Khasi". *Austroasiatic Studies* ed. by Philip N. Jenner, Laurence C. Thompson & Stanley Starosta, Oceanic Linguistics, Special Publication No. 13, 477-522. Honolulu: University of Hawaii.

Jacob, Judith M. 1963. "Prefixation and Infixation in Old Mon, Old Khmer, and Modern Khmer". *Linguistic Comparison in South-East Asia and the Pacific* ed. by Harry L. Shorto, 62-70. London: SOAS.

Jenner, Philip N. and Saveros Pou. 1980-81. "A lexicon of Khmer morphology". *Mon-Khmer Studies* 9-10.1-517.

Kruspe, Nicole D. 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press.

Matisoff, James. 2003. "Aslian: Mon-Khmer of the Malay Peninsula". *Mon-Khmer Studies* 33.1-58.

Peiros, Il'ja J. 2004. *Genetičeskaja klassifikacija avstroaziatskix jazykov*. Moskva: Rossijskij gosudarstvennyj gumanitarnyj universitet (doktorskaja dissertacija).

Radhakrishnan, R. 1981. "The Nancowry word: phonology, affixual morphology and roots of a Nicobarese language". *Current inquiry into language and linguistics*, 37. Edmonton, Alberta: Linguistic Research Inc.

Reid, Lawrence A. 1994. "Morphological evidence for Austric". *Oceanic Linguistics* 33:2.323-344.

Reid, Lawrence A. 1999. "New linguistic evidence for the Austric hypothesis". *Selected Papers from the Eighth International Conference on Austronesian Linguistics* ed. by Elizabeth Zeitoun & Paul Jen-kuei Li, 5-30. Taipei: Academia Sinica.

Reid, Lawrence A. 2005. "The current status of Austric: A review and evaluation of the lexical and morphosyntactic evidence". *The peopling of East Asia: putting together archaeology, linguistics and genetics* ed. by Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas, 81-106. London: Routledge/Curzon.

Schmidt, Pater Wilhelm. 1906. "Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens". *Archiv für Anthropologie, Braunschweig* 5.59-109.

Schmidt, Wilhelm. 1905. "Grundzüge einer Lautlehre der Mon-Khmer-Sprachen". *Denkschrift der Akademie der Wissenschaften, Wien, Philologisch-Historische Klasse* 51.1-233.

Shorto, Harry L. 1963. "The Structural pattern of northern Mon-Khmer languages". *Linguistic Comparison in South-East Asia and the Pacific* ed. by Harry L. Shorto, 45-61. London: SOAS.

Shorto, Harry L. 1971. A dictionary of the Mon inscriptions, from the sixth to the sixteenth centuries, incorporating materials collected by the late C.O. Blagden London Oriental Series 24. London: Oxford University Press.

Shorto, Harry L. 2006. *A Mon-Khmer Comparative Dictionary*. Canberra: Pacific Linguistics.

Sidwell, Paul. 2005. *The Katuic Languages: classification, reconstruction and comparative lexicon*. Munich: Lincom Europa.

Sidwell, Paul. 2006. "Preface to Shorto 2006" *A Mon-Khmer Comparative Dictionary* by Harry L. Shorto, vii-xxv. Canberra: Pacific Linguistics.

Smith, Kenneth, D. 1975. "The velar animal prefix relic in Vietnam languages". *Linguistics of the Tibeto-Burman area* 2:1.1-18.

Solntseva, Nina. 1996. "Case-marked pronouns in the Taoih language". *Mon-Khmer Studies* 26.33-36.

Suwilai Premsrirat. 2002. *Thesaurus of Khmu Dialects in Southeast Asia*. Thailand: Institute of Language and Culture for Rural Development Mahidol University.

Svantesson, Jan-Olof. 1983. "Kammu Phonology and Morphology". *Travaux de l'institut de Linguistique de Lund* 18. Lund: Liber Förlag.

Thomas, David & Robert Headley. 1970. "More on Mon-Khmer subgroupings". *Lingua* 25.398-418.

Thomas, David. 1966. "Mon-Khmer subgroupings in Vietnam". *Studies in Comparative Austroasiatic Linguistics* ed. by Norman H. Zide, 194-202. The Hague: Mouton.