WHY WE DON'T NEED AUSTRIC OR ANY OTHER MACROPHYLA IN SE ASIA: THE SOUTHERN YUNNAN INTERACTION SPHERE

Roger Blench
Kay Williamson Educational Foundation
8, Guest Road, Cambridge CB1 2AL
United Kingdom
Voice/Answerphone 00-44-(0)1223-560687
Mobile 00-44-(0)7967-696804
E-mail R.Blench@odi.org.uk
http://rogerblench.info/RBOP.htm

This printout: January 10, 2010

TABLE OF CONTENTS

TABLES

FIGURES

ACRONYMS

*	regular reconstruction
AD	Anno Domini
BC	Before Christ
BP	Before present
N	nasal
PAA	Proto-Austroasiatic
PAN	Proto-Austronesian
PHM	Proto-Hmong-Mien
PMP	Proto-Malayo-Polynesian
PST	Proto-Sino-Tibetan
V	vowel

1. Introduction

All the language phyla of East Asia have significant common features, including lexical items, vowel systems, semantic and syntactic constructions. This has led scholars to propose genetic connections between individual phyla or else to speculate on very large macrophyla such as Starosta's (2005) PSEA (Proto-South-East Asian). Indeed almost all the major phyla (Sino-Tibetan, Daic [=Tai-Kadai] Hmong-Mien, Austroasiatic and Austronesian) have been connected with one another at different times. Early 'Indo-Chinese' hypotheses linked Daic with Chinese, or later, Sino-Tibetan (Van Driem 2005). Influential for a period was 'Austro-Thai', a hypothesis first advanced by Benedict (1942, 1975), which broadly claimed Austronesian and Daic were related. Benedict (1990) later expanded his view to include Japanese, a direction in which few have followed. Another significant macrophylum proposal is Sino-Austronesian (STAN), propounded by Laurent Sagart (2005a and elsewhere). In the most recent version of this hypothesis, Sagart (2008) proposes that the speakers of STAN were millet farmers.

The fact that these various speculations have yet to resolve into a consensus view should make us wary; many scholars are unable to accept that some types of contact situation make possible loans into even very fundamental vocabulary. Related to this is a failure to resolve significant questions about either the homeland or the antiquity of these phyla. Only Austronesian appears to have a significant consensus; both Taiwan as the homeland and ca. 5500 BP as the period of early expansion (Bellwood 1995). But proposals for Austroasiatic could hardly be more at odds.

Austroasiatic has traditionally been divided into two branches, Mon-Khmer (in SE Asia) and Muṇḍā (in India). Muṇḍā is marked by a dramatic shift of word order compared with the rest of Austroasiatic, a typological feature which has sometimes been allowed to overshadow its otherwise strong similarities with the other branches. But there is strong disagreement between its most well-known researchers concerning the internal structure of the phylum. Diffloth (2005:79) currently considers Austroasiatic to have three primary branches and a complex nested structure, with the earliest dates for diversification placed at 5000 BC[1]. He argues that faunal reconstructions support a 'southern' homeland. A significant challenge to this model has been put forward by Sidwell (2007, 2009), who argues instead for a 'flat' array, in other words, rejecting not only the Muṇḍā/Mon-Khmer split, but all the other proposed internal nodes. Sidwell's latest proposals argue for a Mekong homeland and a much more recent date. Van Driem (2001) canvasses a number of theories, including the 'northern shores of the Bay of Bengal', although this is not based on linguistic evidence. Norman & Mei (1976) and Schuessler 2007[2] have put forward disputed lexical evidence for Austroasiatic loans into Old Chinese this is a fragile foundation for such a major hypothesis.

A hypothesis with a venerable history is Austric, a proposed macrophylum that would unite Austroasiatic and Austronesian, and possibly Daic and Hmong-Mien. The observation that 'Indo-Chinese' might include the languages of the Pacific has an old history; it may have been first advanced by Keane (1882) on linguistic and anthropological grounds. Pater Schmidt (1906) established it in roughly its modern form, and it has had a number subsequent adherents, but remained largely in limbo during much of the twentieth century. Benedict (1976) considered whether Austronesian (or Austro-Thai in his terms) and Austroasiatic could be related but concluded that the observed similarities were due to 'substratumized' Austro-Thai. However, during the 1990s, there was a revival of interest following papers by La Vaughan Hayes (1992, 1997, 1999) who put forward a large number of potential cognates between the two phyla. Although many are not accepted by other scholars, they remain a fruitful basket of suggestions. More influential has been work by Diffloth (1994), Reid (1994, 1999, 2005) and Blust (1996) placing it back into serious consideration. Blust (op. cit) has put forward a scenario for the early expansion and spread of these two phyla, emerging from 'the area in which the Salween, Mekong and Yangzi run parallel at their narrowest watershed'. Austric has also been taken up by archaeologists; Higham (1996, 1998) says quite unambiguously 'the development of rice cultivation in the Yangzi valley took place among people who spoke languages of the Austric phylum'.

[1] Itself somewhat more recent than in previous publications which showed dates as early as 10,000 BC.
[2] Though see sceptical comment in Sagart (2008)

This paper will argue Austric, and indeed all the other macrophyla proposals, are unnecessary; that these similarities result from contact. A credible solution must combine the results of linguistics and archaeology and that whatever combination of contact linguistics and genetic affiliation accounts for the data, it must be within a plausible geographical and chronological framework. It further suggests that the solution is relatively simple; a combination of recent archaeological and linguistic findings now make a coherent narrative. But it depends on two important assumptions;

a) Contact. Contact between two genetically quite unrelated phyla can lead to borrowing of quite fundamental linguistic features
b) Congruence. That historical reconstructions must be congruent with archaeological data. In other words, if crops and livestock are reconstructible to a proto-language then we must assume its speakers were farmers and were located in both space and time in a region where agriculture is archaeologically attested

These assumptions may seem commonplace to Austronesianists working in the Pacific; the interaction of Papuan and Austronesian has long been shown to produce remarkable contact phenomena. Similarly, congruence between the archaeological and linguistic record in Polynesia is an old story. But the situation is not the same in SE Asia, where linguists have been unwilling to let go of genetic explanations for similarities, and have been remarkably insouciant about apparent contradictions between historical linguistics and archaeology[3].This paper will argue the following;

1. Austronesian and Austroasiatic do have significant resemblant vocabulary. Most of this is simply contact in the intermediate periods of their respective expansions, but a few lexical items do appear to reconstruct to the proto-languages of both phyla.
2. Austronesian, Austroasiatic, Hmong-Mien and Daic all show clear evidence of agricultural vocabulary in their proto-languages. They therefore cannot be significantly older than the agriculture attested in the archaeological record and the three mainland phyla must therefore have relatively 'flat' structures.
3. Daic is a branch of Austronesian, either a sister-language to PAN or a parallel branch to PMP. Morphologically, Proto-Daic looked like Austronesian despite its very different appearance today (as languages like Buyang suggest).
4. Daic speakers were in contact with early Austroasiatic and Hmong-Mien speakers in a region between southern Yunnan and northern Vietnam at the period of the earliest attested agriculture, i.e. about 4000 years ago.
5. This accounts for the similarities between these phyla and thus no macrophylum hypotheses need be invoked.

A further historical point must be underlined. Some of the arguments about genetic affiliation seem to be based on the assumption the historical points of contact between Austronesian and Austroasiatic were quite limited; for example, the Chamic presence in Vietnam or the spread of the Malay round the coasts of the SE Asian mainland following the growth of Śrīvijaya. However, archaeology has now gone a long way to counter this assumption. The astonishing similarities between pottery recorded in Thailand, Vietnam and the Philippines (e.g. Solheim 1964, 1992; Yamagata 2008) argue for much more pervasive contact between the mainland and island SE Asia at different points in prehistory and therefore many more opportunities for borrowing.

This paper will focus on vocabulary and prefix morphology. Many other similarities and features common to SE Asian language phyla have been noted (Enfield 2003, 2005). It is assumed that the historical explanation for these will be along the same lines as the scenario sketched out here.

[3] For example, it has recently been proposed that the Hoabhinian (>18,000 BP) might correlate with Austroasiatic, a phylum for which all researchers agree there are solid reconstructions pointing to agriculture.

2. Some of that common vocabulary

The presence of common lexemes between apparently unrelated families at a deep level in SE Asia underlies much of the discussion about genetic affiliation[4]. A good example of this is the word for 'bird' (Table 1).

Table 1. A common lexeme for 'bird'

Language	Form
Proto-Malayo-Polynesian	*manuk
Buyang	ma-nuk[11]
Proto-Tai	*n-lok
Proto-Hmong-Mien	*m-nok
Chinese	niǎo (鳥) [?]

Discussed in Ostapirat (2005:118). This term for 'bird' is PMP and has been retained in Daic in almost its original form, at least in Buyang and Lakkia. PAN for 'bird' is something like **qayam* which becomes 'chicken' in Austronesian daughter languages, but no Formosan cognates of **manuk* have been recorded. Schuessler (2007:401) suggests the innovative term in Chinese may be cognate but this is definitely open to doubt.

Even more striking in its prevalence is the word for 'eye' (Table 2);

Table 2. A common lexeme for 'eye'

Language	Form
PAN	*maCa
Bunun	*mataʔ
PAA	*kmat
Proto-Daic	*mata
Buyang	ma-ta[54]
PHM	*mɻɛyH

The prefix in the PAA form is based on forms such as Khasi *khmat* 'eye', but the most common form in many branches of Austroasiatic is simply *mat* and this is most likely to be inherited from an original Daic loanword.

Table 4 shows a common lexeme for 'hair';

Table 3. A common lexeme for 'hair'

Language	Form
PMP	*buSék
Kavalan	bukes *hair of the head*
PAA	*suk; *suuk; *suək; *sək
Old Khmer	suk
Bahnar	sɔk
Proto-Muṇḍā	*sok
Chrau	sənɔːʔ *body hair*
Khasi	shñiuh
proto-Kra	*m-səm
proto-Hlai	*h-nom
proto-Tai	*phom

Discussed in Benedict (1976) and Shorto (2006: #467) who considered the connection with Austronesian 'doubtful' although he gives no reason. There is every reason to think that the Austronesian and Austroasiatic forms are cognate; the Austronesian prefix is deleted in Austroasiatic. Moreover, Austroasiatic

[4] I would like to acknowledge an unpublished paper on this subject by Waruno Mahdi (ined.) presented at the 11 ICAL in Aussois, June 2009, from which a number of citations are taken.

is cognate with the PMP form (mistakenly given as PAN in Zorc (1995)) as Formosan forms all have the metathesis *bukeS*. However, the more puzzling aspect of this is whether the Daic terms are cognate. They suggest that the Kra forms are closest to the original proto-Daic. Some of the Austroasiatic witnesses point to an original with a nasal in medial position; in which case PAAS would have been more like **sənɔ(ɔ)k*. This makes the cognacy of the Daic forms more likely though not certain.

Table 4 shows a common lexeme for 'bone' in all four phyla;

Table 4. A common lexeme for 'bone'

Language	Form
PAN	*Cuqəlaŋ
Paiwan	tsuqelał
Tagbwana	tuʔlaŋ
Moken	kelan
Betsileo	tólaŋə
PAA	*ʤəʔaŋ
Khmu	ʧəʔaŋ
Temiar	tulag
Gelao	taŋ (D2)
PHM	*ʧuŋ(X)

Gelao is exceptional in retaining the old root, since Kra languages seem have generally replaced it with the [?] unrelated **dak*. But Austroasiatic retains clear traces of a CVCVC structure, so the morphology of its source lexeme may have resembled PMP forms, reflected in Tagbwana, where the glottal stop would be the eroded segment attested in Taiwan. The Moken form is interesting because it has deleted the initial syllable in the same way as many Austronesian-Daic descendants. As a consequence, the proto-Daic form may have been longer, perhaps also **Culaŋ*, in which form it was borrowed into Austroasiatic. Hmong-Mien may have subsequently borrowed from Daic. Temiar *tulag* looks suspiciously like the Austronesian forms and may be a more recent borrowing.

The interchange of 'bow/arrow/shoot' is attested elsewhere in the world and does not seem an unlikely semantic shift. When the shift to 'crossbow' occurred is less certain; the earliest archaeological record of a crossbow is a bronze crossbow mechanism dating to around 600 BC from a grave burial at Qufu (Zhu Fenghan *Ancient Chinese Bronzes* p. 274). The first reliable textual record of crossbow usage is in the battle of Ma-Ling, Lingyi, China in 341 BC[5]. Within a century, the crossbow was well developed and widely used in China. Table 5 shows a common lexeme for 'bow/arrow/crossbow/shoot';

Table 5. A common lexeme for 'bow/arrow/crossbow/shoot'

Phylum	Branch	Language	shoot	bow	crossbow
Austronesian		PAN		*panaq	
	Formosan	Ami		panáq	
	Formosan	Tsou	pono		
	Philippines	Ilokano		pana	
Austroasiatic	Vietic	Proto-Vietic			*s-naaʔ
	Vietic	Thavung			thanə̂:
	Vietic	Mương	baŋ[3]		
	Khmeric	Khmer	paŋ បាញ់	theaʔnuʔ ធនុ	snaa ស្នា
	Pearic	Pear [Kompong Thom]	ṃpɔŋ (clf.)		thna
	Bahnaric	Sre	paŋ		
	Bahnaric	Laven	peŋ		hnaa
	Bahnaric	Stieng	peːŋ		səna:
	Bahnaric	Sedang		ʂãŋ	
	Katuic	Kuy	mpan	tnùu	nha:, sna:

[5] Curiously at almost the same time (400 BC), the first records of the *gastraphetes* appear in Hellenic Greece.

Phylum	Branch	Language	shoot	bow	crossbow
	Khmuic	Khmu	paŋ		
	Mangic	Bolyu		tɛuŋ^{53}	
	Palaungic	Riang [Sak]	pəŋ^{1}		
	Monic	Mon	paŋ	pɔn ʊʂ̊	sɑn
	Monic	Old Mon		tŋaʔ	
	Monic	Nyah Kur	péŋ		thnùu
	Nicobarese	Nancowry	hafə̀ŋ		fə̀ŋ
Daic	Kra-Dai	Proto-Southern Kra-Dai		*hɲɯ:	
Daic	Kra-Dai	Laha		na^{132}	
Daic	Hlai	Proto-Hlai		*hɲɯ:	
Daic	Hlai	Hlai (Li)	tseɯ^{53}		
Daic	Kam-Sui	Sui		hna$_{B1}$	
Hmong-Mien	Mienic	P-Mienic			*hnək^{D}
Sino-Tibetan	Sinitic	Old Chinese			*nâʔ (弩)
	Nungish	Nung			thəna
	Nungish	T'rung			tānā
	Na	Naxi			tana

The only Austroasiatic languages without these roots are the Aslian and Khasian branches. Aslian speakers switched to blowpipes and hence seem to have adopted a quite distinctive vocabulary. Khasian (and probably Muṇḍā) seem also have left the crossbow behind.

Working out the complex history of this root is quite difficult. Clearly there is an old Austronesian root **pana(q)* reflected in proto-Austroasiatic **paŋ*. This has not obvious reflexes in Daic, which seems to have borrowed from Austroasiatic or Hmong-Mien forms with initial *hn-*. This might be connected with a switch from bows to crossbows. Austroasiatic languages developed a *t^{h}V-* prefix for 'crossbow', which is conceivably connected with forms such as Hlai *tseɯ*^{53} 'to shoot' and could be a compound meaning something like 'shooting bow'. The *t^{h}V-* prefix is then reduced along various pathways, to t→s→h→ø. The velar in Mienic points to a borrowing from a language with a velar nasal such as Sedang, although these forms are not particularly close. Sinitic forms for 'crossbow' look as they have a different source from other Sino-Tibetan languages. Nungish and Na have clearly borrowed 'crossbow' from Austroasiatic, with the original borrowing into Nung which retains the initial *t^{h}V-*. The exact source is somewhat mysterious, since the closest languages, Palaungic and Mangic do not have these forms, and it is hard to imagine the geographical frame for contact with Vietic or Khmer. It seems likely that the reduced forms in Sino-Tibetan such as Naxi *tana* are internal Sino-Tibetan borrowings. Daic sources are somewhat defective lexically, so we cannot be sure about the absence of some items. However, the forms for 'bow' appear to be late borrowings from Austroasiatic languages such as Bahnaric or from Mienic. Laha na^{132} may be a borrowing from Sinitic rather than a reduced form of other Daic.

An intriguing case, because it is a food plant and thus evidence for subsistence, is the word for 'sesame'. Table 6 shows the reflexes of *#ləŋa:* 'sesame' in SE Asian languages;

Table 6. Reflexes of *#ləŋa:* 'sesame' in SE Asian languages

Phylum	Branch	Language	Attestation	Comment	Source
Austronesian	Philippines	Tagalog	liŋa		
Austronesian	Chamic	PC	*laŋa		Thurgood (1999)
Austronesian	Malayic	Malay	leŋa		
Austroasiatic	Palaungic	Palaung	ləŋa		Milne (1931)
Austroasiatic	Palaungic	P-Waic	*rŋaʔ		Diffloth (1984)
Austroasiatic	Monic	Middle Mon	laṅau		
Austroasiatic	Khmeric	Khmer	lɔŋɔ:		Vidal et al. (1969)
Austroasiatic	Katuic	Pacoh	laŋe:		Watson (n.d.)
Austroasiatic	Bahnaric	PSB	*rəŋa, *ləŋa		Sidwell (2000)
Austroasiatic	Khmuic	Khabit	ləŋaa		K & S (1999)
Daic	Kra	Buyang	ŋaa		Ostapirat (2000)
Daic	Kam	Sui	ʔŋaa		Ostapirat (2000)
Daic	Tai	Thai	ˈŋaa งา		SEAlang
Hmong-Mien	Mien	Mun of Hainan	taa nyim		Shintani & Yang (1990)
Sino-Tibetan	Loloish	Mpi	nɤŋ[2]		Bradley (1979a)
Sino-Tibetan	Loloish	Bisu	hnám		Bradley (1979a)

The widespread presence of this root also presents a historical problem as sesame would not normally be considered sufficiently old in the region to be attested in this way. The original homeland of sesame is the subject of some debate, since it is grown from Africa to China and has been found in many excavations in Near Eastern sites (Blench 2003). Earlier authors (e.g. Nayar & Mehra 1970) saw West Africa as its homeland, since most of the wild relatives of sesame are found there. However, during the 1980s, Bedigian *et al.* (1985) also Bedigian (2003) proposed that its progenitor was the Indian *Sesamum orientale* var. *malabaricum* which today grows wild on granitic outcrops and is found in a weedy form all over the subcontinent. More recently, Hiremath & Patil (1999) have advanced a strong case for *S. mulayanum*, also occurring in India. Archaeological evidence for sesame in ancient India is sparse; *Sesamum* is present during the Mature Harappan period at Mohenjo-Daro, 2600/2500-2000 cal. BC (Fuller & Madella 2001). Although a single radiocarbon date like this is normally treated with scepticism by archaeologists, its antiquity would have to be of this order to reach Taiwan and become a credible candidate for PAN.

Another intriguing shared lexeme is the word for 'dog', *#-tʃɔ:*, apparently shared between Austroasiatic and Austronesian at a rather fundamental level. Table 7 shows the reflexes of this root;

Table 7. Reflexes of *#atʃɔ:k* 'dog' in SE Asian languages

Phylum	Branch	Language	Attestation	Source
Austronesian		PAN	*asu/wasu	Blust (2002)
Austronesian	Formosan	Thao	atu	Blust (2003)
Austronesian	Formosan	Pazeh	wazu	
Austronesian	Malayic	PMP	*asu	
Austronesian	Chamic	PC	*ʔasɔw	Thurgood (1999)
Austroasiatic	Palaungic	Lawa	sɔʔ	
Austroasiatic	Vietic	Ruc	ʔacɔ:[3]	Alves (200x)
Austroasiatic	Monic	Proto-Monic	*clur	Diffloth (1984)
Austroasiatic	Khmeric	Angkorian Khmer	ca	
Austroasiatic	Pearic	Pear	tʃɔ:k	Headley (1977)
Austroasiatic	Bahnaric	PSB	*sɔ:	Sidwell (2000)
Austroasiatic	Katuic	PK	*ʔəcɑ:	Sidwell (2000)
Austroasiatic	Khmuic	P-Khmuic	*sɔʔ	Premsirat (2002)
Austroasiatic	Aslian	Semelai	cɔɔh	
Austroasiatic	Khasian	War Jaintia H	kʰsu	Brightbill et al. (2007)
Austroasiatic	Munḍā	Kharia	sɔ[-lɔʔ]	Stampe
Daic	Tai	Thai	sù waan สุวาน	SEAlang

Phylum	Branch	Language	Attestation	Source
Daic	Tai	Lao	cɔ: จ๋	Kerr (1972)
Sino-Tibetan	Loloish	Phu Ka	tʃɯ^{35}	Edmondson (n.d.)
Sino-Tibetan	Tibetic	rGyalthang	tshə̄	Krisadawan (2000)
Sino-Tibetan	Naga	Garo	acak	Burling (2003)

The term is so pervasive in Austronesian and Austroasiatic it would be difficult to argue that it does not go back to the proto-language in both cases. Unfortunately for the broader argument of this paper, Daic languages in China have all replaced reflexes of a possible Austronesian inherited term with **maa* which is probably a borrowing from Hmong-Mien (PHM **hmaŋ*^{C}). The attested cases of this root in Daic, such as in Lao, must be more recent borrowings from Austroasiatic. Sagart (2008:143) points out that a Southern Yuè word for 'dog', recorded in the *Shuō Wén*, a Chinese character dictionary first published in 100 AD, is pronounced *ou-sou* or *ou-ʂou*.

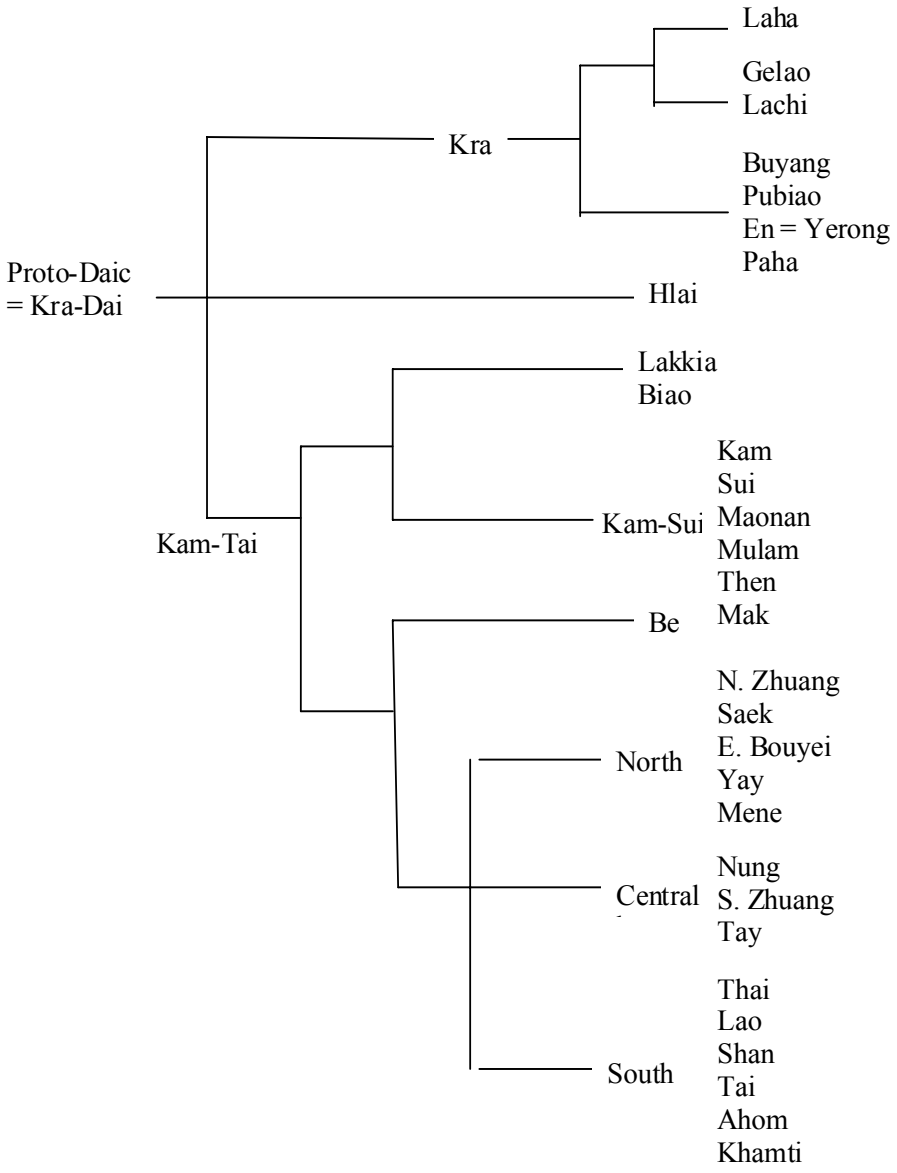
Examples could be multiplied but these are some of the 'best' in the sense of being attested very widely; many of the other words cited (for example in Reid 2005) have a suspiciously patchy distribution. Also interesting are the apparently shared morphological elements (Reid 1994, 1999). Much hangs on the estimates of individual linguists as to the likelihood of morphological borrowing. But many of the affixes are far from pervasive paradigms and may be the result of lexical borrowing and re-analysis, of which there are many examples.

3. The Daic link to Austronesian

A problem for Benedict's (1942, 1975) 'Austro-Thai', the hypothesis which broadly claimed Austronesian and Daic were related, was that Daic and Austronesian appear to be so very different on the surface; Daic is highly tonal with very short words, Austronesian is non-tonal and tends to have CVCV stems plus affixes. Hence the tendency was to treat Daic as isolated or to link it with Sino-Tibetan, which appears much more similar in terms of morphology. Benedict also explained some of the apparent similarities between Austroasiatic and Daic as 'substratumised' Austro-Thai, an argument which Diffloth (1977) effectively demolished. Thurgood (1994) argued that the apparent relationship with Austronesian is simply that of loanwords. However, Ostapirat (2005) demonstrates regular sound-correspondences in a way more acceptable to mainstream comparativists in support of a genetic affiliation.

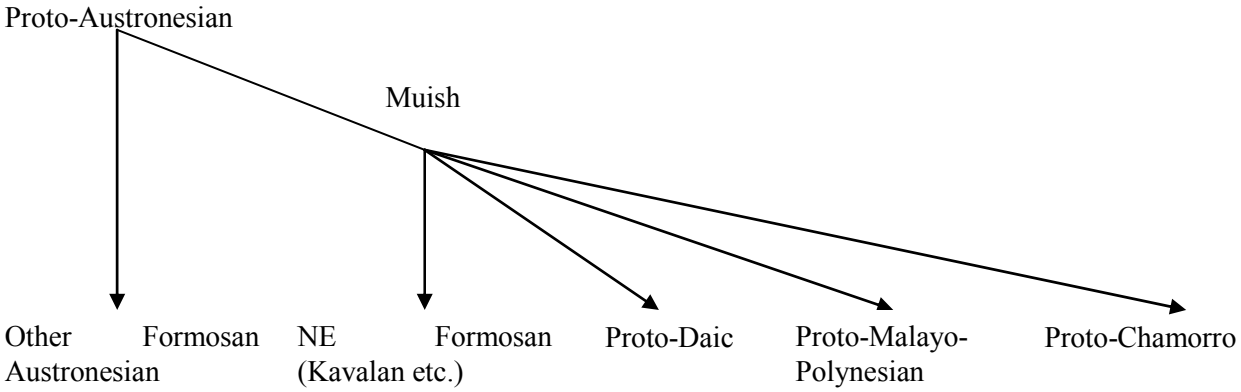
The Daic or Tai-Kadai languages are spoken from southern Thailand into Laos, Cambodia and China. Overviews of the phylum are given in Edmondson & Solnit (1988, 1997a) and Diller et al. (2008). Figure 1 shows the internal classification of Daic updated from Edmondson & Solnit (1997b). Ostapirat (2005) presents a rather different view with five primary branches, splitting Be, Tai and Kam-Sui, but this is not a consensus view nor is it supported with lexical or phonological evidence. Figure 1 shows the internal classification of Daic updated from Edmondson & Solnit (1997b).

Figure 1. Internal classification of Daic



Ostapirat links his 'proto-Kra-Dai' with the Austronesian reconstructions of Blust (e.g. 1999) and Zorc (1995) but the evidence for the place of Daic within Austronesian remains unresolved. Sagart (2004, 2005b) puts Daic on a level corresponding to Malayo-Polynesian as branch of 'Muish', part of his proposed phylogeny of Formosan Austronesian. Figure 2 shows the ancestry of Daic according to Sagart (2005b);

Figure 2. Ancestry of Daic according to Sagart (2005b)



Source: Condensed and adapted from Sagart (2005b)

To support the idea that Daic is a sister-language to PMP, Sagart (2004) cites evidence from Buyang, a mainland Daic language, showing conservation of typical Austronesian morphology (Table 8).

Table 8. Evidence for Daic-Austronesian links

Gloss	Buyang	PAn	Malayo-Polynesian
die	ma-tɛ^{54}	maCay	matay
eye	ma-ta^{54}	maCa	mata
bird	ma-nuk^{11}		manuk
head	qa-ɗ'u^{312}	quluh	quluh
louse	qa-tu^{54}	kuCu	kutu
fart	qa-tut^{54}	qetut	
raw	qa-ʔdip^{54}	qudip	
cover v.	ta-qup^{11}		WMP ta(ŋ)kup

Source: adapted from Sagart (2004)

This neatly demonstrates that typical Austronesian morphology was retained by Daic *after* the arrival of speakers back on the mainland and that the reduced forms now typical of most Daic languages are a later development. Some key items like 'bird' **manuk* and 'nose' **ijuŋ* are found only in PMP, and not on Taiwan, which does support the view of Sagart.

If this linguistic scenario is accepted, then proto-Daic speakers would have migrated back to the mainland from the southern tip of Taiwan about 4000 BP. At this period, the Chinese mainland would have presented a very different ethnolinguistic picture from today. The main body of the Chinese population would have been further north and there would have been a diverse body of minority ethnic groups, speaking Hmong-Mien, Austroasiatic and other Sino-Tibetan languages (of which Tujia and Bai may well be the only remnants today) as well as entirely lost language phyla. The speakers of Daic would have spread inland slowly, gradually diversifying. Probably their most ancient branches would been assimilated by the southward expansion of the Han in all the areas near the coast[6]. However, most importantly, they would have encountered the early speakers of Austroasiatic and (probably) Hmong-Mien at the point when these language phyla were just initiating their expansion. It would have been at this point that contact would have occurred, hence the surprising cognates between Austronesian and Austroasiatic[7]. The location of this interaction would have been in the region of Southern Yunnan and northern Vietnam, here christened the South Yunnan Interaction Sphere (SYIS). At this point, characteristic Austronesian morphology would have

[6] Luo (1997) points to an interesting borrowing, probably from proto-Tai into Old Chinese, the word for 'rainbow'. Reconstructed to proto-Tai as **Druŋ*, it is cited in the earliest Zhou lexicon, the Erh Ya as *dì dòng* and reconstructed by Pulleyblank as *təwŋ'*.

[7] Although expressed in very different language, this is broadly the conclusion which Benedict (1976) reached with his explanation of and Austro-Thai 'substratum'.

persisted in Daic languages, since the pervasive prefix deletion and tonal evolution had made only a limited impact on their structure.

Such a scenario is only credible if the expansion of Austroasiatic and Hmong-Mien are recent. This is not problematic to argue in the case of Hmong-Mien, which is not very diverse internally; however, Austroasiatic is often thought to be very internally differentiated. The next section will argue that the transparent reconstruction of agriculture in Daic, Austroasiatic and Hmong-Mien implies that we must assume relatively recent dates for the diversification of these, in order for there to be congruence with the archaeology.

4. The reconstruction of agriculture in proto-languages

4.1 Introduction

The possibility that the expansions of many of the world's language phyla were driven by agriculture has had considerable airtime recently (e.g. the contributors to Bellwood & Renfrew 2002). The likelihood of this scenario can be questioned, but logically, if crops and livestock terms are reconstructible in a proto-language, it is reasonable to assume its speakers were familiar with agriculture, hence the requirement for congruence with the archaeological record. This section summarises the evidence for the reconstruction of crop and livestock names in Daic, Austroasiatic and Hmong-Mien.

4.2 Daic

Both crops and domestic animals can be reconstructed for proto-Daic. Ostapirat (2000) presents some glosses that appear to be shared across all three branches, including 'pig', 'dog' and at least some crops. Table 9 shows these reconstructible items;

Table 9. Daic lexicon illustrative of subsistence

Language	chicken	pig	dog	sesame	'yam'
Gelao	qai	map	mpau	ŋklau	mbø
Lachi	kɛ	mye	m	—	mɦa
Laha	kəi	məu	maa	—	mal
Paha	qai	muu	maa	ŋaa	man
Buyang	ʔai	muu	—	ŋaa	man
Biao	qai	m̥uu	m̥aa	ŋɦɯɑ	mɦən
Hlai	khai	pou	pou	keɯ	man
Sui	qaai	m̥uu	m̥aa	ʔŋaa	man
Tai	kai	muu	maa	ŋaa	man

Source: Ostapirat (2000)

The word for 'yam' is not easy to interpret, since this root is applied to taro in a number of languages (Burmese *mun*, Vietnamese *môn*). It is possible it is a texture of borrowings in Daic and not a reconstructible root. Blench (2005) has presented some evidence for thinking that speakers of proto-Daic were not originally rice-growers, and that they borrowed cultivation techniques from Austroasiatic speakers. Reconstruction has yet to produce positive evidence for their subsistence strategies, and it may be that they were originally cultivators of tubers such as taro. In support of this is Ostapirat's (2005:119) comparison between PAN **biRaq* (for the cultivated *Alocasia* sp.) and the Daic forms;

Table 10. Taro in Austronesian and Daic

Language	Attestation
PAN	*biRaq
Laha	haak D2
Paha	pɣaak D2
Buyang	ðaak D2
Hlai	geek
Kam-Sui	ʔɣaak
Tai	phɯak

Source: compiled from Ostapirat (2000, 2005)

It seems credible that the Austronesian name for *Alocasia* sp. was transferred to *Colocasia* on the mainland. However, this is unrelated to the Austroasiatic root, which must represent a separate domestication occurrence.

In an interesting addendum to the proto-Tai reconstructions of Li (1977), Luo (1997) points to the large number of proto-Tai terms associated with rice agriculture. For example;

Table 11. Rice production terms in proto-Tai

Gloss	proto-Tai
to slash, clear land	*thraaŋ
to release water	*khaaŋ
to ear (crops)	*maan
young grilled rice	*hmaw
chaff of unhusked rice	*kaak
barnyard grass	*hwaŋ

Source: Luo (1997)

4.3 Austroasiatic

The controversy over the dating and internal classification of Austroasiatic is summarised in §1. However, indirect evidence does seem to point to a relatively shallow time-depth for Austroasiatic. Agricultural terminology appears to unambiguously reconstruct to the proto-language. Table 12 shows crop reconstructions in Austroasiatic crops and their approximate incidence across individual branches[8].

Table 12. Crop reconstructions in Austroasiatic

Gloss	Reconstruction	Comment
rice (general)	#ɓa:ʔ	Found in seven branches
rice-grain	*sŋɔ:ʔ	Reconstructs only to Proto-Mon-Khmer
paddy rice	#srɔ	Found in three branches including Munda
husked rice	#rəkau	Found in seven branches including Munda
foxtail millet	#səŋkɔɔy	Found in seven branches
taro	#trawʔ	all branches except Aslian
sesame	#ləŋa	Found in six branches
banana	#tVlVy	Found in six branches
betel pepper	#mpluw	Found in six branches

The best attested crop is taro, for which a common root is attested almost everywhere. Rice is similarly widespread, and includes Muṇḍā, which points strongly to its presence in the earliest period. These forms are consistent with the claim by Diffloth (2005) that Austroasiatic speakers typically spread along river valleys in the early period of their expansion, seeking swampy ground to cultivate taro. But they are not congruent with a date of 7000 BP. There is evidence for the rapid expansion of the Neolithic in the Yunnan/Northern Vietnam borderland, for example at Baiyuncun and Phung Nguyen some 4000 years ago

[8] Full datasets are included in Blench (forthcoming b)

(Higham 2002: 85 ff.). These sites are characterised by the 'incised and impressed' pottery that spreads very rapidly across the region in this period (Rispoli 2008). If agriculture itself is ca. 4200 BP, the initial dispersal of proto-Austroasiatic cannot logically be earlier than this. If this is the case, then Austroasiatic is unlikely to have an intricate nested structure, because this would not allow sufficient time for such a structure to develop. Sidwell's 'flat array' model of Austroasiatic is thus more plausible.

4.4 Hmong-Mien

A broadly similar argument applies to Hmong-Mien; it has a simple internal structure and agriculture can be reconstructed to the proto-language. Although the reconstruction of agricultural vocabulary is clear, none of the terms are transparently related to the other two phyla (except for the word for 'dog' mentioned above). Borrowings from Old Chinese dominate proto-Hmong-Mien, which anyway has many fruit-crops and other plants typical of a drier climate, not generally characteristic of the SE Asian region. Table 13 shows proposals for the Hmong-Mien subsistence lexicon;

Table 13. Proposals for the Hmong-Mien subsistence lexicon

Item	Reconstruction	Source
bean	*dup	< Chinese
buckwheat	*ɟæu	cf. Chinese
chicken	*Kəi	< Chinese
cucumber	*K^{w}a	< Chinese
eggplant	*ɟa	cf. Chinese
pear	*rəy	< Chinese
plum	*hlɟəŋX	
rice, cooked	*hnrə̂aŋH	
rice, husked	*tuX	< Chinese
rice plant	*mbləu	
taro	*wouH	< Chinese
buffalo	*ŋiuŋ	< Chinese
dog	*qluwX	
duck	*ʔap	< Chinese
sheep/goat	*yuŋ	< Chinese

Adapted from Ratliff (in press)

The dates of the primary expansion of proto-Hmong-Mien are likely to be in the same time-period and the similarities with Daic and Austroasiatic, documented in the Tables in §2., are a result of the same period of intensive interaction.

5. Conclusions: further research

This paper seeks and explanation for the apparent similarities between Austroasiatic and Austronesian given that almost any hypothesis places a significant geographical distance between their homelands. It accepts that Daic is a branch of Austronesian and that its earliest speakers may have left Taiwan during the period of the earliest Austronesian maritime expansion which also resulted in the Malayo-Polynesian languages. It furthermore suggests that early Daic would have looked structurally very like Austronesian, which accounts for the synchronic similarities with forms in other language phyla. It also assumes that in situations of intense bilingualism, fundamental vocabulary can easily be borrowed, a proposition that should be evident form the numerous borrowings from Old Chinese into proto-Hmong-Mien.

Figure 3 shows a map which illustrates schematically the proposed South Yunnan Interaction Sphere and illustrates how Austronesian roots could have been transferred to Austroasiatic and Hmong-Mien without invoking macrophylum hypotheses.

Figure 3. Map showing the location of the proposed South Yunnan interaction sphere
Error! Objects cannot be created from editing field codes.

References

Bedigian, Dorothea & Jack R. Harlan 1986. Evidence for the cultivation of sesame in the ancient world. *Economic Botany* 40, 1986:137-154.

Bellwood, Peter & C. Renfrew (eds.) 2002. *Examining the Farming/ Language Dispersal Hypothesis*. (McDonald Institute Monographs.) Cambridge: McDonald Institute for Archaeological Research.

Bellwood, Peter 1995. Austronesian prehistory in Southeast Asia: homeland, expansion and transformation. In *The Austronesians: historical and comparative perspectives*. 96-111. Peter Bellwood, James J. Fox and Darrell Tryon eds. Canberra: ANU Press.

Benedict, Paul King 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *American Anthropologist* 44:576-601.

Benedict, Paul King 1975. *Austro-Thai Language and Culture, with a glossary of roots*. New Haven: Human Relations Area Files Press.

Benedict, Paul King 1976. Austro-Thai and Austroasiatic. *Oceanic Linguistics Special Publications, No. 13, Austroasiatic Studies Part I*. 1-36. Hawai'i: University of Hawai'i Press.

Benedict, Paul King 1990. *Japanese/Austro-Tai*. (Linguistica Extranea Studia 20). Ann Arbor: Karoma Publishers.

Blench, Roger M. 2003. The movement of cultivated plants between Africa and India in prehistory. In: K. Neumann, A. Butler & S. Kahlhaber (eds.) *Food, fuel and fields: progress in African Archaeobotany*. 273-292. Köln: Heinrich-Barth-Institut.

Blust, Robert A. 1996. Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology. In: Ward H. Goodenough, ed., *Prehistoric Settlement of the Pacific* (Transactions of the American Philosophical Society, Vol. 86, Pt. 5). 117-160. Philadelphia: American Philosophical Society.

Blust, Robert A. 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In: *Selected papers from Eighth International Conference on Austronesian linguistics*. E. Zeitoun & P. Jen-Kuei Li eds. 31-94. Taipei: Academica Sinica.

Diffloth, G. 1977. Mon-Khmer initial palatals and "substratumized" Austro-Thai. *Mon-Khmer Studies*, VI:29-57.

Diffloth, Gérard 1994. The lexical evidence for Austric, so far. *Oceanic Linguistics,* 33:309-22.

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.

Diller, Anthony V.N., Jerold A. Edmondson & Luo Yongxian (eds). 2008. *The Tai-Kadai Languages*. Routledge: Abington, Oxon UK.

Edmondson, J.A. and D.B. Solnit eds. 1988. *Comparative Kadai: linguistic studies beyond Tai*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Edmondson, J.A. and D.B. Solnit eds. 1997a. *Comparative Kadai: the Tai branch*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Edmondson, J.A. and D.B. Solnit eds. 1997b. Introduction. In: *Comparative Kadai: the Tai branch*. 1-32. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Enfield, N.J. 2003. *Linguistic epidemiology: semantics and grammar of language contact in mainland Southeast Asia*. London: Routledge Curzon.

Enfield, N.J. 2005. Areal Linguistics and Mainland Southeast Asia. *Annual review of Anthropology*, 34: 181-206.

Fuller, D.Q. and M. Madella 2001. Issues in Harappan archaeobotany: retrospect and prospect. In: S. Settar & R. Korisettar (eds.) *Indian archaeology in retrospect, II. Protohistory.* New Delhi 2001: 317-390 (Oxford and IBH).

Hayes, L.H. 1992. On the Track of Austric: Part I. *Mon-Khmer Studies,* xxx.

Hayes, L.H. 1997. On the Track of Austric: Part II. Consonant mutation in early Austroasiatic', *Mon-Khmer Studies,* 27:13-44.

Hayes, L.H. 1999. On the track of Austric: Part III. Basic vocabulary comparison', *Mon-Khmer Studies,* 29:1-34.

Higham, Charles 1996. Archaeology and linguistics in Southeast Asia: implications of the Austric hypothesis. In: *Bulletin of the Indo-Pacific Prehistory Association*, 14: 110-118.

Higham, Charles 1998. Archaeology, linguistics and the expansion of the Southeast Asian Neolithic. In: *Archaeology and Language II*, R.M. Blench and M. Spriggs eds. 103-114. London: Routledge.

Higham, Charles 2002. *Early cultures of mainland Southeast Asia*. Bangkok: River Books.
Hiremath, S.C. & C.G. Patil 1999. Genome homology and putative progenitor of Sesame (*Sesamum indicum* L.). *Journal of Cytological Genetics* 34(1):69-74.
Luo, Yongxian 1997. Expanding the proto-Tai lexicon - a supplement to Li (1977). *Mon-Khmer Studies*, 27:271-298.
Mahdi, Waruno 2009. *Some obscure Austroasiatic borrowings in Indonesian and Old Malay*. Paper circulated at the 11th ICAL, Aussois.
Moore, Elizabeth H. 2007. *Early landscapes of Myanmar*. Bangkok: River Books.
Nayar, M.N. & K.L. Mehra 1970. Sesame: its uses, botany, cytogenetics and origin. *Economic Botany* 24,1, 1970:20-31.
Norman, J. and Mei, Tsu-lin 1976. The Austroasiatics in ancient South China: some lexical evidence. *Monumenta Serica* 32: 274-301.
Ostapirat, Weera 2000. Proto-Kra. *Linguistics of the Tibeto-Burman area*, 23.1.
Ostapirat, Weera 2005. Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics,* 107-131. London: Routledge Curzon.
Ratliff, Martha in press. *Hmong-Mien language history*. Pacific Linguistics. Canberra.
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. in: E. Zeitoun and P. J.-K. Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics,* Taipei: Academia Sinica
Reid, Lawrence A. 2005. The current status of Austric. in: Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. 132-160. London: Routledge Curzon.
Rispoli, F. 2008. The incised and impressed pottery of Mainland Southeast Asia: following the paths of Neolithization. *East and West*, 57:235-304.
Sagart, Laurent 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics,* 43 (2): 411-444.
Sagart, Laurent 2005a. Sino-Tibetan-Austronesian: an updated and improved argument. in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics.* 161-176. London: Routledge Curzon.
Sagart, Laurent 2005b. Tai-Kadai as a subgroup of Austronesian. in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics.* 177-181. London: Routledge Curzon.
Sagart, Laurent 2008. The expansion of Setaria farmers in East Asia: a linguistic and archaeological model. In: Alicia Sanchez-Mazas, Blench, R.M., Ross, M.D., I. Peiros & Marie Lin eds. *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics.* 133-158. London: Routledge.
Schmidt, Pater Wilhelm 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. *Archiv für Anthropologie*, 5:59-109. Braunschweig,
Schuessler, Axel 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: Hawai'i' University Press.
Sidwell, Paul 2007. *Comparative Mon-Khmer Linguistics in the 20th Century: Where From, Where To*? Presentation at SEALS.
Sidwell, Paul 2009. *The Austroasiatic Central Riverine Hypothesis*. Keynote address, SEALS, XIX.
Solheim, W.G. 1964. Further relationships of the Sa Huynh – Kalanay Pottery Tradition. *Asian Perspectives*, 8(1):196-211.
Solheim, W.G. 1992. Nusantao traders beyond Southeast Asia. In: Glover, I., Pornchai Suchitta & J. Villiers eds. *Early metallurgy, trade and urban centers in Thailand and Southeast Asia*. 199-225. Bangkok: White Lotus.
Starosta, Stanley 2005. Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific. in: Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas eds. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics.* 182-197. London: Routledge Curzon.
Thurgood, Graham 1994. Tai-Kadai and Austronesian: the nature of the relationship. *Oceanic Linguistics,* 33(2):345-368.

Van Driem, G. L. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region containing an Introduction to the Symbiotic Theory of Language*. 2 vols. Leiden: Brill.

Van Driem, G. L. 2005. Tibeto-Burman vs. Indo-Chinese: implications for population geneticists, archaeologists and prehistorians. In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. 81-106. London: Routledge Curzon.

Yamagata, Mariko 2008. *Archaeological research on the prehistoric interrelations beyond the South China Sea.* Waseda University.

Zorc, R. D. 1995. A glossary of Austronesian reconstructions. In: *Comparative Austronesian dictionary. Part I. Fascicule 2*. D. Tryon et al. eds. 1105-1197. Berlin/New York: Mouton de Gruyter.