7 *The role of agriculture in the evolution of mainland Southeast Asian language phyla*

ROGER BLENCH

1 Introduction

The emergence of mainland Southeast Asian (MSEA) nations from decades of war and the gradual opening up of individual countries has created a new focus on the region. Southeast Asia presents an intriguing mix, combining highly diverse ethnolinguistic groups with generally small populations, and more numerous peoples, such as the Thai, Burmese, Lao, Vietnamese and Khmer, who cover significant swathes of territory and are politically dominant. Historically the vast majority of languages of Southeast Asia were unwritten and remain poorly described, but the existence of scripts has created something of a focus on texts. Recent years have seen the emergence of much new data, often quite difficult to access; nonetheless, the linguistic map is gradually becoming clearer, although many questions remain.

There are five major language phyla in mainland Southeast Asia, Austroasiatic, Austronesian, Daic, Sino-Tibetan and Hmong-Mien and no true isolates, except on offshore islands. Despite a considerable expansion of research in recent years, models for the dates, homelands and engines of expansion of these phyla are markedly absent from the literature, as are convincing correlations with archaeological and genetic research. A claim that has generated considerable discussion in recent years is the importance of agriculture and thus demographic expansion in accounting for ethnolinguistic geography. The chapter considers whether the reconstruction of agricultural terminology in individual language phyla supports this claim, and if so, what can be said about the dating of individual phyla. Since Austroasiatic is discussed at greater length elsewhere in this volume (for example, Diffloth, Sidwell and Blench) its treatment will be abbreviated here.

The five language phyla that dominate MSEA are:

- Sino-Tibetan
- Hmong-Mien [=Miao-Yao]
- Austroasiatic [partly = Mon-Khmer]
- Austronesian
- Daic [=Tai-Kadai, Kra-Dai]

N. J. Enfield, editor.
Dynamics of Human Diversity, 125-152. Pacific Linguistics, 2011

There are virtually no isolates, except Andamanese, possibly Shompen (Blench 2007, in press) and the extinct Kenaboi (Hajek 1998). The protectiveness of the Indian government has made it difficult to establish clearly the nature of Andamanese and even the relation of the languages to one another. Abbi (2006) has gone some way to remedying this, but still the type of rich lexical and grammatical data which could underlie hypotheses of long-range connections remains elusive. A lack of credible archaeology has meant that there are no dates for first settlement of the archipelago. The Andamanese may have been there for a very long time, or they may been brought there in the boats of others in the last few thousand years. Other foragers, notably the Orang Asli of the Malay peninsula and the negritos of the Philippines, now speak Austroasiatic and Austronesian languages respectively. Linguistic reconstruction does suggest a substrate vocabulary in their modern-day speech (Reid 1994a, b), but we are far from being able to link this fragmentary lexicon with other language phyla.

Due to long periods of interaction and extensive multilingualism, MSEA language phyla have developed many convergent characteristics, as well as being extensively relexified from dominant or contact languages (Enfield 2003, this volume). As a consequence, numerous macrophyla hypotheses have been advanced linking together almost any combination of phyla, notably Austric (for example Higham 1996; Reid 2005 and references therein), Daic/Austronesian (Schlegel 1901; Ostapirat 2005), Austro-Thai (Benedict 1942, 1975), Sino-Tibetan-Austronesian (STAN) (Sagart 2005a, 2008) and 'Proto-East Asian' (Starosta 2005; van Driem 2008). Despite this, our understanding of the proto-lexicon, dating and patterns of dispersal from the homelands of individual phyla remains both sketchy and controversial. In the case of Sino-Tibetan, a failure to make available comparative materials that purportedly underlie proposed reconstructions has made assessment of the true situation difficult.

In Africa and Oceania there is a relatively long tradition of combining archaeology and linguistics to develop a synthesis of prehistory. This may partly reflect the absence of historical documents and large-scale polities, but it is also a consequence of intellectual traditions which favour interdisciplinarity. With a few exceptions, such an integrated prehistory remains to be created for MSEA. It is, however, clearly needed, if we are to go beyond epigraphy to a more global account of the movements and evolution of present-day populations. The complex patterns of language phyla must correlate with the archaeology of the region in some fashion. Historical linguistics can provide both a general insight into the peopling of the region and also help research both the transition from foraging to farming and the history of individual crops and livestock species. This chapter is intended to present preliminary results of research into the reconstructibility of the agricultural lexicon in the language phyla of Southeast Asia and suggest their interpretation in the light of recent archaeological data. The main hypotheses relating to the peopling of Southeast Asia and the inception of agriculture are outlined, and then each of the major language phyla is reviewed in turn. The conclusions draw together the evidence for the dating and possible expansion of these phyla; it should be emphasised that this is highly preliminary. Many of the speculations presented here will need considerably more work, linguistic and archaeological, to refine their application to archaeological data.

2 Archaeology of Mainland Southeast Asia

The origins of the current populations of Southeast Asia have been the subject of much debate. On the basis that Australoid populations must have passed through the region, a 'two-layers' model has generally been proposed. Broadly speaking, this assumes that there

was an original peopling of Australoids, phenotypically similar to modern Andamanese and other negritos, and that these were replaced by mongoloids, apparently migrating down from present-day China, although not at that point Sinitic-speakers. However, skeletal and mtDNA evidence for this hypothesis has been conspicuous by its absence, although well-preserved finds of the appropriate date are very rare. Matsumura et al. (this volume) report on skeletal material dated to 10,450 +/- 300 years BP at Hang Cho, in Luong Son district, southwest of Hanoi, which they claim shows negrito affinities. Lewis et al. (2008) describe a terminal Pleistocene cremation burial on the island of Palawan dated to 9500–9000 BP, the earliest yet known in the region. The languages of these earlier foragers must remain unknown, apart from Andamanese and possible relic vocabulary in Aslian languages.

Southeast Asian societies today are clearly very agriculturally-based, with rice production a fundamental activity throughout the region. Despite this, there is little or no direct evidence for the transition to agriculture, and even the date at which this took place is disputed (Bellwood 1997). Indeed, Southeast Asian archaeology shows a very distinctive 'Holocene gap' with a marked lack of sites between ca. 8–4000 BP, the reasons for which are unknown (Joyce C. White, personal communication). One notable exception is the archaeological site of Da But, in Thanh Hoa province of Vietnam, an early Neolithic cemetery and shell midden, radiocarbon dated to 5085 BC (Vinh 1991). The subsistence strategies of the Da But people were mixed hunting, gathering, and fishing, but evidence for animal husbandry and paddy rice cultivation remains controversial.

The most widespread claim for the dating of the Neolithic transition is that of Higham (2004:47) who notes 'The pattern of intrusive agriculturalists settling inland valleys in southern China, while the coast continued to be occupied by affluent foraging groups, is repeated in the Red River area and the contiguous coast of Vietnam'. The type-site for this type of agriculture is Phung Nguyen, about 200 km inland from Halong Bay. Dates remain problematic, but the adjacent site of Co Loa has been dated to 2000 BC (Lai Van Toi 1999). In summarising the situation, Higham says:

> We find agricultural settlements being founded in the lower Red River valley, along the course of the Mekong and its tributaries, and in the Chao Phraya valley…The dates for initial settlement, as far as they are known, are approximately the same with none earlier than about 2300 BC. Most intriguingly, the pottery vessels in many of the sites over a broad area have a similar mode of decoration. The sites reveal extended inhumation graves and an economy incorporating rice cultivation and the raising of domestic stock. (Higham 2002:352).

In contrast, White (1995) and White et al. (2004:123) say 'based primarily on dates from basal deposits from Ban Chiang and Ban Tong, two long-term settlements in the Kumphawapi catchment, that societies cultivating plants appeared in the region by the mid-fourth millennium BC cal.' If so, agriculture would be as much as 1500 years earlier than the Higham model.

Higham and Higham (2009) are now proposing a new chronology for the beginning of the Neolithic in MSEA, based on the new C14 chronology of Ban Non Wat, which derives from a large sample of directly dated shell, analysed with Bayesian statistics. This would imply a revision of all existing dates towards the more recent period, from 1800/1700 to 1100 BC. (Fiorella Rispoli, personal communication) considers the recent C14 dated excavations at An Son and Da Kai in South Vietnam are fully in accord with this. On the basis of comparisons between the sites she has excavated in Central Thailand and most of

the other Neolithic sites in Thailand, Vietnam and Yunnan, the new dates 'put all the tiles in the right place', linking MSEA with Yunnan as well as Guangxi/Guangdong.

The claims of White and her collaborators are based on indirect environmental evidence, rather than direct archaeobotanical materials. The conflict between their views and the main body of Southeast Asian archaeologists might be reconciled if farming was preceded by a long period of intensive landscape management, but distinguishing between these two interpretations may be difficult based purely on the archaeological record.

3 Sino-Tibetan

The Sino-Tibetan phylum has more speakers than any other language phylum, largely due to the inclusion of the Sinitic branch, composed of the many varieties of Chinese. Despite some two centuries of study and publication, the subclassification of Sino-Tibetan remains highly controversial, as does its external affiliation (Blench 2008). Considering the importance of Sino-Tibetan and its history of scholarship, there is a striking lack of agreement as to its internal classification. Some key questions are:

- whether the primary branching is Sinitic (that is all Chinese languages) versus the remainder (usually called Tibeto-Burman) or whether Sinitic is simply part of one branch, for example Bodic et cetera. Certainly the distinctiveness of Sinitic is far from proven.
- what are the inter-relations of its branches?
- can it be linked with other phyla such as Austronesian or Caucasian (as proposed for example by Sagart 2005a, 2005b; Starostin 1991).

Broadly speaking, the opposing camps are those who consider Sinitic as the primary branching of Sino-Tibetan (Wolfenden 1927; Benedict 1972, 1976; Bodman 1980; Weidert 1987; Matisoff 2003, 2008; Bradley 1997; Thurgood and La Polla 2003) and those who situate it within the remaining languages, consequently applying the name Tibeto-Burman to the whole phylum (Shafer 1966/67; van Driem 1997). Sinitic would thus be incorporated within the group conventionally defined in opposition to it. The two markedly different views are shown in Figures 1 and 2.

The groups represented in Figure 1 are by and large 'geographic' categories; Kamarupan and Himalayan have no status as linguistic subgroupings. Even this view has never been justified in print, despite the space afforded by the 800 pages of Matisoff (2003). Moreover, since Matisoff excludes many small branches of Sino-Tibetan and joins many others at a single node, this is far from a fully worked-out theory. An agnostic alternative is represented by van Driem (2005), in his 'fallen leaves' schema (Figure 2). van Driem's model presents no assumptions at all about subgrouping except to map already well-recognised groups. This is a geographical model, which places generally agreed subgroups in proximity, with area of the ellipse corresponding to their size, but advances no hypothesis about their ultimate relationships. Whether this represents progress is debatable, but the 'fallen leaves' model has the virtue of treating all branches of Sino-Tibetan as of equal status and requiring that their position be ultimately defined. Van Driem would argue that this is a fair representation of the current state of our knowledge.

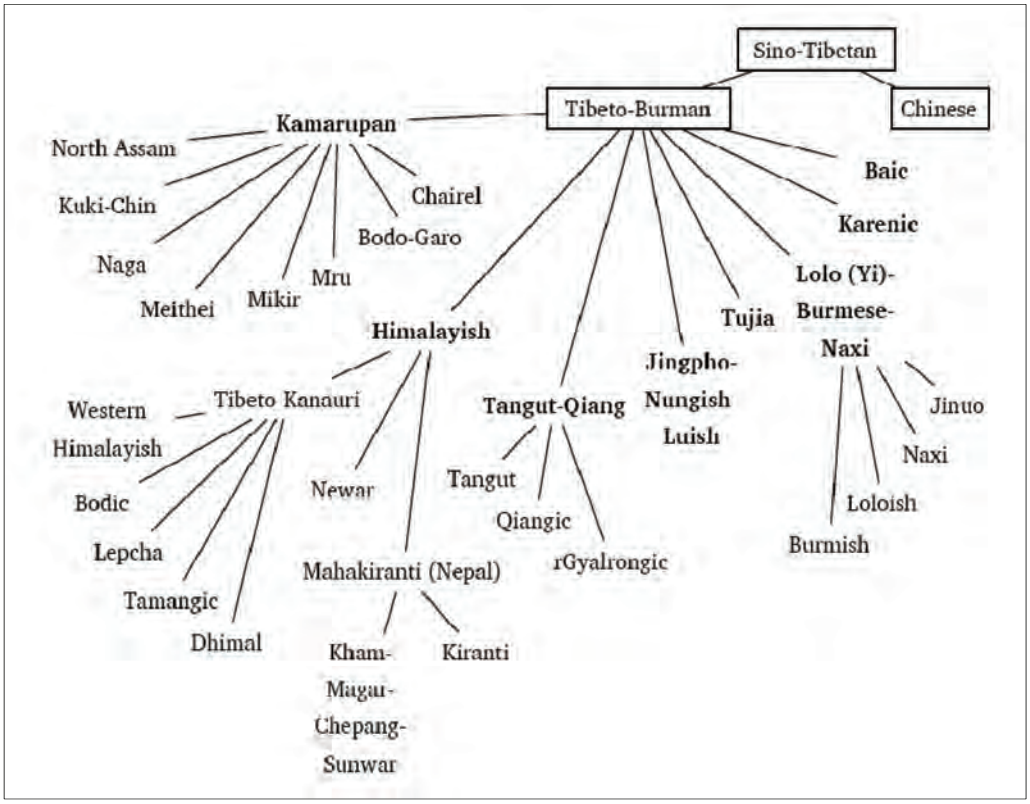

Figure 1: Sino-Tibetan according to Matisoff (2008).

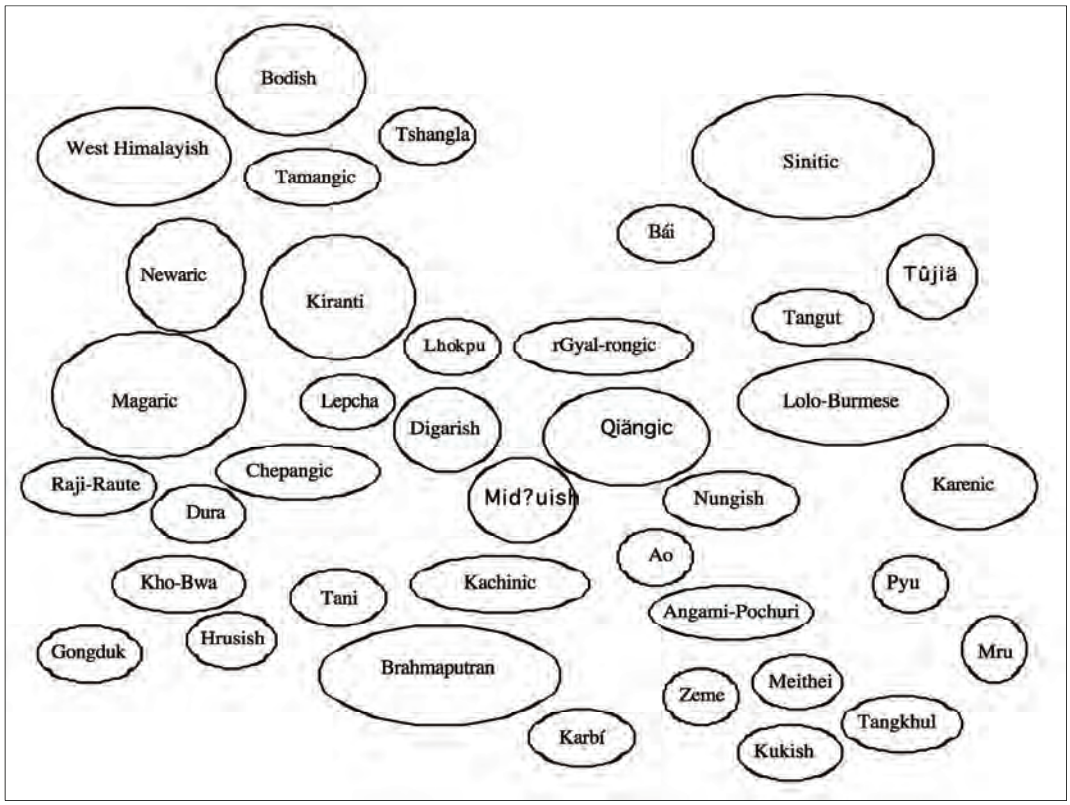

Figure 2: 'Fallen leaves' model of Sino-Tibetan according to van Driem (many places).

Both these classifications essentially show large parallel arrays, with van Driem being the extreme version of the agnostic view. Clearly, the building of a hierarchical model of Sino-Tibetan appears to be a long way off. Neither model seems to give fair weight to the highly diverse languages of Northeast India, for which documentation is gradually becoming available[1]. In terms of internal diversity, the region from the southern flanks of the Himalayas to the Assam region is massively diverse synchronically, with large numbers of small subgroups which appear to be very different from one another. We may have to suppose an original diversity more characteristic of Northeast Asia or parts of the Amazon. As Northeast India and adjacent regions open up, the striking linguistic diversity of this region is becoming more apparent. For example, Sherdukpen, Bugun (Dondrup 1990) Lishpa, and Butpa are listed in the Ethnologue as Tibetic languages with no evidence. Examination of the actual data on these languages provides almost no support for such an affiliation; indeed even their membership of Sino-Tibetan is only supported by a small number of lexemes which could well be borrowings. Similarly with Hrusish spoken between Assam and Bhutan in Arunachal Pradesh[2] whose classification remains highly uncertain (Shafer 1947; Simon 1970). It has yet to be proven that some groups are Sino-Tibetan at all rather than isolates with a Sino-Tibetan superstrate.

Related to this diversity is the absence of the classic grain-based agriculture implicit in much of the Sino-Tibetan literature. The supposedly Tani-affiliated Milang emphasise vegeculture and hunting strongly in contrast to cereals, and terms for domestic animals, for example, all appear to be recent borrowings. The Sulung, who also speak a language of uncertain affiliation (Tayeng 1990) base their subsistence on sago-exploitation and hunting. Either we assume that the classification of these languages is in error, or that it is not the case that we can confidently reconstruct *any* agricultural terms to Proto Sino-Tibetan, simply because there are no certain attestations in numerous subgroups, especially in the Himalayan and Northeast Indian branches. Logically, therefore, the region of Arunachal Pradesh may well be a remaining zone of high diversity in early Sino-Tibetan from which the later, secondary rice cultivating, pig-producing cultures evolved. Unfortunately, the absence of well-dated, stratified archaeological sites in this region makes it impossible to correlate with archaeology at present.

Northeast India is far from the only problem with Sino-Tibetan. The Sinitic region also includes at least two languages, Tujia and Bai, which are single branches of Sino-Tibetan and which seem to have no particular relationship with Sinitic except for numerous borrowings at many historical levels. Unlike the foragers of Northeast India, these groups are strongly agricultural, but appear to preserve archaic vocabulary pointing to a pre-Sino-Tibetan presence in the region. Tujia has a raft of non-Sinitic agricultural terminology, which has either been innovated or shows links with other Sino-Tibetan subgroups (Table 1).

---

1 Thanks especially to Mark Post, who has kindly collected and made available a wealth of local publications on Northeast India.

2 Shafer (1947:184) says ‘A glance the Hruso vocabulary of any author except Campbell reveals almost no similarity to any known Tibeto-Burmic language, and it has been with the greatest difficulty that enough Hruso comparisons have been gathered to show the genetic relationship of that language to Tibeto-Burmic and to establish a few correspondences’.

Table 1: Unexpected agricultural vocabulary in Tujia

Tujia	**Gloss**	**Parallels**
si^{1} li^{1}	rice	
ye^{3}	cooked rice	
zi^{3}	cooked rice	
qie^{1} hhe^{1}	glutinous rice	cf. Qiang *qhəˈ*
oŋ^{1}ba^{1}	sorghum	
loŋ^{1} moŋ^{4}	barley	
si^{3} tuo^{4}	garlic	
kuo^{1} su^{1}	ginger	
ge^{3}	horse	
wu^{2}	cow	? reduction of widespread #*ŋu*
zi^{4}	pig	? cf. Chinese *shǐ* (豕)
ruo^{2}	goat	
ha^{4} lie^{3}	dog	
ra^{3}	chicken	cf. Dayang (Qiangic) *ro*
sa^{4}	duck	? Austroasiatic e.g. Bugan *mtʃa*^{13}
ŋo^{3}	goose	cf. Mantsi Meo Vac (Loloish) *ŋɔ*^{31} *ŋɔ*^{44}

Bai shows many fewer such items, probably because so much of its basic lexicon has been replaced by Chinese loans (Table 2).

Table 2: Unexpected agricultural vocabulary in Bai

Bai	**Gloss**	**Parallels**
te^{42}	pig	cf. Kayah Li thé, Biao Min twə^{4}
ky^{21}	buckwheat	
me^{55} zo^{42}	barley	cf. proto-Lolo-Burmese *zəy²
tɕũ^{33} ŋɔ^{21}	sheep	
χua^{55} lao^{31}	cat	cf. Naxi χua lɛ

These data suggest a pre-Sinitic presence of Sino-Tibetan-speaking agriculturalists throughout much of this region that was largely absorbed following the expansion of the Han Chinese. These *might* be correlated with the earliest Neolithic communities in North China such as the Péilígǎng or Císhān (6500 BP onwards) but Hmong-Mien speakers are equally likely candidates. Fuller et al. (2008) have recently questioned the dating of many of the early rice-producing communities in central China, suggesting that many finds are wild rice and that domestication only really begins by 6500 BP.

Given this situation, we cannot confidently reconstruct *any* agricultural terms to Proto-Sino-Tibetan, simply because there are no certain attestations in numerous subgroups, especially in the Himalayan and Northeast Indian branches. The presence of a term in Lahu unfortunately does not guarantee its reconstructibility to Proto-Sino-Tibetan. Agriculture presumably developed well after the primary dispersal of Sino-Tibetan, which must therefore be considerably earlier than the other language phyla in the region. Given this, there are widespread terms attesting agriculture in well-studied branches such as Sinitic, Karenic and Lolo-Burmese and these must certainly reflect the importance of farming in the secondary expansion of Sino-Tibetan. Table 3 presents my proposals for quasi-

reconstructions (that is, not the result of inspection of systematic sound-correspondences) of crop and livestock terms.

Table 3: Widely attested agricultural terms in Sino-Tibetan

	Gloss	**Comment**
#mei	rice	[also in Daic and Hmong-Mien]
#ʃan	rice	[also in Daic]
#tʃɔk	foxtail millet	[also in Mienic and ?Austronesian]
#ŋwV	cow, ox	[also in Daic and Austroasiatic]
#brak	pig	[also in Austronesian]

It is notable that all these terms are found outside Sino-Tibetan, especially in Daic, which points strongly to a period of intense interaction in the early phases of the intensification of agriculture.

The evidence for early Sino-Tibetan is marked by gaps in the data; an absence of reflexes for agricultural terms in many of the smaller branches of Sino-Tibetan, a lack of evidence for coherent internal structures and a failure of congruence with archaeology and genetics. Given this, any hypothesis concerning its spread and diversification must be speculative and subject to revision. However, we can do better than any of the claims presently on the table by presenting an account which at least does not contradict the interdisciplinary evidence. With this in mind, the following scenario is put forward as a model of the development of the phylum:

- The earliest speakers of Sino-Tibetan were highly diverse foragers living in an arc between the slopes of the Himalayas and Assam/Arunachal Pradesh up to 10,000 years ago.
- Some spoke early Sino-Tibetan languages, others unknown languages now present only as substrates unless Kusunda is a relic of this period.
- Seasonal foragers exploited the high Tibetan Plateau from 7500 BP.
- By perhaps 6–5000 BP a ‘livestock revolution’ took place. Yak herders moved up and settled the Tibetan Plateau permanently. Pigs were domesticated in China among non-Sino-Tibetan speakers.
- Foragers in Northeast India probably began to practise vegeculture (taro, plantains) and arboriculture (sago) and animal management (mithun) by 6000 BP.
- By 5000 BP diverse early Sino-Tibetan groups spread eastwards to China. Sinitic is not a primary branch but simply one of many migratory groups.
- Proto-Tujia, proto-Bai and probably others met unknown populations (Hmong-Mienic? pre-Austronesians?) with domestic pigs, while also cultivating and beginning to domesticate rice.
- Proto-Sinitic speakers encountered early Altaic speakers with foxtail millet and other crops.
- The Sinitic languages expanded southwards, assimilating or encapsulating many small groups. They encountered Hmong-Mien speakers with rice and switch millet terminology to rice.

- Cold zone cereals (buckwheat, foxtail and Panicum millets) and perhaps also taro were moved from gathering to domestication in the montane areas on the fringes of the Himalayas.
- Rice moved up from India but also westwards from China (hence hybridised types) and overlays older cereals where ecologically possible.
- Ruminants (cows, sheep, goats) spread downwards into China from Central Asia 4400 BP (? Altaic for small ruminants but not cattle).
- Tibetic speakers undergwent a major expansion (when?) assimilating linguistic diversity on the Plateau.
- Rice invaded the lowland vegecultural zones rather later, pushing taro into residual systems.
- Groups such as early Burmic spread southwards, fragmenting Austroasiatic-speaking peoples.

Figure 3 shows a highly simplified map of the early phases of these movements.

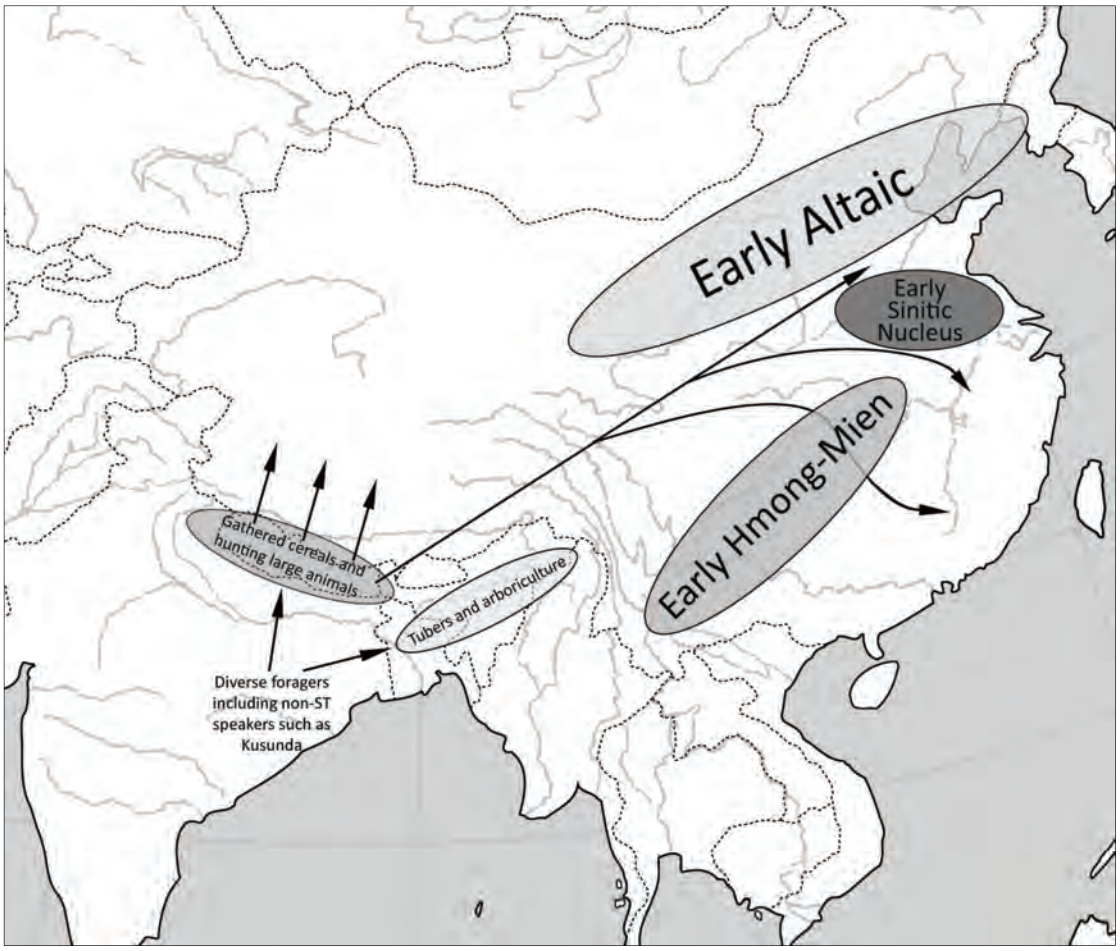

Figure 3: A possible model of early Sino-Tibetan expansion.

4 Hmong-Mien

The Hmong-Mien [=Miao-Yao] languages are spoken mostly in China with some groups also in Laos, Vietnam and Thailand (Niederer 1998, 2004; Ratliff). Their centre of gravity is between the Yangzi and the Mekong rivers and the extensions southwards may be as recent as the last few centuries. Hmong-Mien languages are quite closely related to one another, and although the Ethnologue lists some 32 languages, many of these are mutually intelligible lects. The linguistic geography of Hmong-Mien speakers suggests strongly that they were scattered by the incoming Han and probably forced southwards in the last 3–2000 years. Many agricultural terms can be reconstructed to proto-Hmong-Mien but most of them are either Chinese borrowings or resemble Chinese closely. It seems likely that pre-Hmong-Mien speakers have a long history in China, and that they can be linked with early Neolithic cultures, but Chinese cultural dominance has made this hard to detect. Figure 4 shows a 'tree' of Hmong-Mien languages in Niederer (2004).

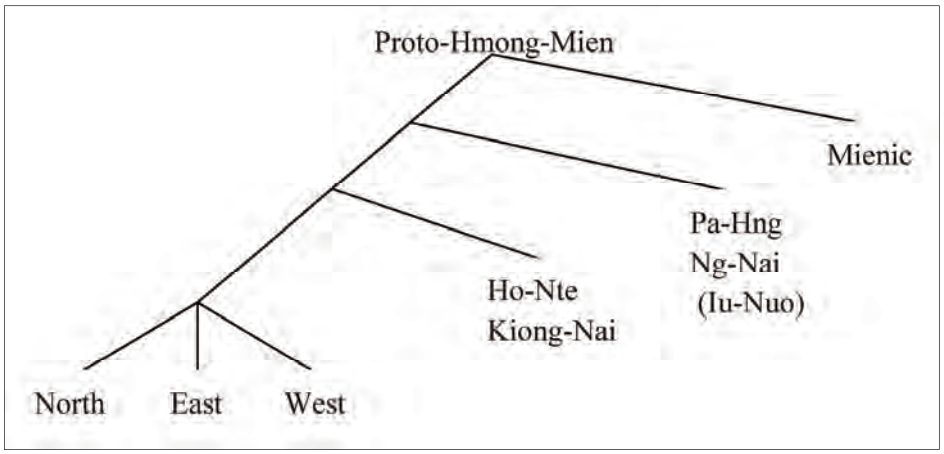

Figure 4: Classification of the Hmong-Mien languages according to Niederer (2004).

The Hmong-Mien proto-language is likely to be older than is apparent from the lexical data. The reconstruction of agricultural vocabulary is clear, but so are borrowings from Old Chinese into the proto-language. Moreover, proto-Hmong-Mien has many fruit-crops and other plants typical of a drier climate which are not generally characteristic of the Southeast Asian region. Table 4 shows proposals for the Hmong-Mien subsistence lexicon adapted from Ratliff.

Table 4: Proposals for the Hmong-Mien subsistence lexicon (adapted from Ratliff)

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	*hlj̊əŋX	
rice, cooked	*hnrəaŋH	

Item	Reconstruction	Source
rice, husked	*tuX	< Chinese
rice plant	*mbləu	
taro	*wouH	< Chinese
buffalo	*ŋiuŋ	< Chinese
dog	*qluwX	
duck	*ʔap	< Chinese
sheep/goat	*yuŋ	< Chinese

Although Ratliff assumes 'buckwheat' is a borrowing from Chinese, this is apparently historically unlikely (Laurent Sagart, personal communication) and it is probable it was adopted into Sinitic. Similarly, 'eggplant' resembles terms in many languages of the region and is probably not a borrowing from Sinitic. Whatever the final resolution of the various etymological debates, proto-Hmong-Mien as presently understood is too late to be identified with the earliest agricultural sites in the Yangzi and other regions of Central China. However, it is not unlikely that speakers of pre-proto-Hmong-Mien were present in this zone.

5 Austroasiatic

Austroasiatic languages remain the most poorly researched of all those in the region. Many are not documented at all and some recently reported in China are still not classified with certainty. Although there have been many proposals, there are no proto-Austroasiatic reconstructions with published justifications. The nearest approach to this is the 'Mon-Khmer' etymological dictionary of Shorto (2006), which identifies a large number of common roots attested across several branches. However, few of them are pan-Austroasiatic and may reflect regional, local or subgroup innovations. It is therefore as yet unclear whether, for example, we can draw inferences concerning the environment in the homeland of Austroasiatic, as has been claimed (Diffloth 2005, this volume). Sidwell and Blench (this volume) review some of the theories concerning the classification and antiquity of Austroasiatic and argue for a relatively recent dispersal along the Mekong basin, marked by incised and impressed pottery (Rispoli 2008; cf. White, this volume). These arguments will not be repeated here and this section will concentrate on the reconstruction of agricultural vocabulary. Van Driem (this volume) argues almost the exact opposite, for an early domestication of rice in Northeast India and an expansion from west to east. Regrettably, the failure to cite any concrete linguistic data supporting his argument makes it hard to evaluate.

Indirect evidence points to a relatively shallow time-depth for Austroasiatic, since a broad variety of agricultural terminology can be reconstructed to the proto-language. The most well-known crop is taro (*Colocasia esculenta*), for which a common root is attested almost everywhere (Table 5). Reflexes of *#traw?* occur throughout Austroasiatic, and Shorto (2006:475) reconstructs a form *$*t_2$raw?* for his proto-Mon-Khmer. Table 5 shows the distribution of reflexes of *#traw?*.

Although the model presented by Sidwell and Blench for the structure and dispersal of Austroasiatic is quite different, it is consistent with the claim by Diffloth (2005) that Austroasiatic speakers typically spread along river valleys in the early period of their expansion seeking waterlogged soils suitable for taro. Rice terminology is widespread, and includes Muṇḍā, but is not attested in as many branches as taro and therefore should not be

treated as proto-Austroasiatic. Table 6 shows quasi-reconstructions of Austroasiatic names for crops and the numbers of individual branches for which reflexes are attested[3].

This distribution suggests that almost all other crops were adopted subsequent to the dispersal of Austroasiatic and that after taro, hill-rice and foxtail millet were key crops. Paddy-rice was apparently quite late despite its dominance in agricultural systems today. Ferlus (1996) makes the intriguing suggestion that there has been a *glissement sémantique* [semantic slippage] between taro and rice terms, presumably via the concept of 'staple crop' (though see Diffloth, this volume, for a sceptical response). This underlines the relative antiquity of taro in Austroasiatic subsistence systems. If agriculture itself is ca. 4300 BP, the initial dispersal of proto-Austroasiatic would not be earlier than this. If this is the case, then Austroasiatic is unlikely to have an intricate nested structure, because the time would be insufficient for such a structure to develop.

Table 5: Reflexes of *#trawʔ* ,'taro' in Austroasiatic

Branch	Language	Attestation	Gloss in source
Palaungic	Riang	sroʔ	
Palaungic	Palaung	tɔh	
Palaungic	Danaw	kăro[1]	
Palaungic	Proto-Wa	kroʔ	
Palaungic	Lamet	ruəʔ	
Palaungic	Khang	hɔ	
Monic	Mon	krao	
Monic	Nyah Kur	traw	
Vietic	Thavung	tʰoo[3]	
Vietic	Vietnamese	sọ	
Vietic	Proto-Vietic	*sroˀ	
Khmeric	Old Khmer	trav	
Khmeric	Khmer	traːv	
Khmuic	Khmu	sroʔ	
Pearic	Chong	kʰreː[A]	
Bahnaric	PSB	*təraw	
Bahnaric	East Bahnar	trɔɔu	amaranth
Katuic	PK	*craw	
Katuic	Bru	ʔarạw	
Katuic	Kuy	ʔaaràaw	
Katuic	Sre	traw	
Katuic	Mlabri	kwaaj	
Katuic	Ong	raw	
Khasian	Khasi	shriew	arum
Muṇḍā	Sora	saro	*Caladium esculentum*
Muṇḍā	Mundari	saɽu	edible root
Muṇḍā	Santal	saru	

Apart from crops, the speakers of proto-Austroasiatic were also enthusiastic livestock producers. Table 7 shows that almost all the major species found in the region today were already known to speakers at an early period, except horse, donkey and sheep.

3 Full datasets are included in Blench (forthcoming)

The most surprising of these is the goat, which is poorly attested archaeologically but for which the linguistic evidence is very strong. It is also notable that aquatic-adapted poultry, such as geese and ducks, appear to be older than chickens.

This evidence is consistent with a relatively late date for the dispersal of Austroasiatic, which seems to have spread rapidly over a large region, as the 'flat' structure of the phylum suggests. Given the importance of taro and other aquatic terminology, it is reasonable to link this with the wide distribution of 'incised and zone-impressed' pottery 'across parts of far southern China, northern Vietnam and Thailand after about 2500 BC' (Bellwood 2005:132; Rispoli 2008; cf. White, this volume). Rice, millet and chickens would have been adopted midway through the expansion, and paddy rice would have come to replace taro as the principal subsistence crop relatively late.

Table 6: 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 Muṇḍā
husked rice	#rəkau	Found in seven branches including Muṇḍā
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

Table 7: Livestock reconstructions in Austroasiatic

Gloss	Reconstruction	Comment
bovid	#ŋwV	Widespread but does not necessarily apply to domestic species
cow	#[rə]mɔɔk	Found in six branches including Muṇḍā
buffalo	#krəpaaw	Found in all branches excluding Muṇḍā
buffalo	#triik	Found in six branches with possible Muṇḍā cognate
pig	#kliik	Found in six branches
pig	#kruul	Found in six branches
goat	#bɛɛŋ	Found in ten branches with doubtful Muṇḍā cognate
dog	#atʃɔ:k	All branches
cat	#miaw	Found in eight branches
chicken	#syiar	Found in six branches
goose/duck	#ŋaŋ	Found in nine branches but referent varies
duck	#tʃaa[k]	Found in nine branches

6 Austronesian

One of the most persuasive narratives in recent prehistory has been that of the Austronesian expansion. Deriving from the original hypothesis of the kinship of over a thousand languages in Southeast Asia and the Pacific, it was first given its modern form by Dempwolff (1938). However, Dempwolff omitted to clearly identify and situate the

languages of the indigenous peoples of Taiwan, an omission rectified by the second major figure in Austronesian studies, Isidore Dyen (1963). Blust (1976, 1995, 1999) may have been the first author to clearly establish that the diversity of Formosan languages required them to be ancestral to all others and to constitute an array of primary branches. This hypothesis was adopted by Peter Bellwood (1984/5) to model the archaeological evidence, whence emerged a story about the ancestors of the Austronesians leaving Taiwan by means of highly developed sailing technology and reaching the furthest shores of the Pacific as well as the coast of East Africa. A Neolithic package was deemed to accompany these ocean navigators, consisting of pigs, dogs and chickens, rice, pottery and stone adzes, as well as distinctive types of jewellery, such as the nephrite *linglingo*. Various sub-narratives such as the 'express-train to Polynesia' (Diamond 2001) reached high-profile journals and the idea has acquired a certain currency in global prehistory. Blust's hierarchy of nodes branching from the Austronesian tree up to Oceanic, the branch identified with the Lapita potters and ultimately giving rise to Polynesian, seemed to correlate with this early expansion.

The Austronesian migration has further developed into a more general narrative about demographic expansion in prehistory which has it that the spread of many of the world's language phyla driven by agriculture (Bellwood 2005; Bellwood and Renfrew 2002). This model has always had its detractors (Meacham 1984/1985; Solheim 1984/1985; Oppenheimer 2004; Szabó and O'Connor 2004; Terrell 2004) but their striking failure to engage with the linguistic evidence has tended to undermine the substance of their arguments. Moreover, in some areas, notably Near Oceania and Polynesia, it would be hard to deny demographic expansion, since this was the colonisation of previously unoccupied territory. Nonetheless, in recent years there has been a rising chorus of discontent; archaeologists are increasingly claiming that the data doesn't fit a simple demic expansion model. Linguists have been less vocal, but then the number of linguists interested in bigger picture of Austronesian is quite restricted. But with Denham (2004), Donohue and Denham (2010), Blench (2005, 2009, 2010), Lewis et al. (2008) and Bulbeck (2008) the chorus of discontent is too loud to be ignored. The claim, put simply, is that assemblages seem to be rather complex and not to correspond to a simple model of incoming Neolithic farmers replacing foragers. Moreover, some of the key elements in the proposed Austronesian subsistence package are simply not turning up in excavations. The patterns of material culture in prehistory seem to point to earlier and more complex inter-island interactions than the Austronesian expansion model would imply.

One of the key building blocks of the Austronesian expansion hypothesis has been the apparent reconstructibility of key economic terms, both for domestic animals and crops. If the findings (or absences) in the archaeological record are to be taken at face value, then there are problems with these reconstructions. There is a dichotomy between animals and plants, since vegeculture systems could have carried domesticated species across the Austronesian world prior to the expansion from Taiwan, whereas this cannot be the case for animals unless they are attested in the archaeozoological record. There is a specific point concerning pigs, dogs and chickens[4]. It has been shown that the majority of pigs in island Southeast Asia originate not from Taiwan, but from the mainland, probably Vietnam (Hongo et al. 2002). Moreover, they are conspicuously absent from the archaeological record in the main islands until significantly later than the Austronesian expansion (Dobney et al. 2008). There is a small pocket of domestic pig in assemblages in Taiwan,

4 Thanks to Phil Piper for discussions on this point.

and the extreme northern Philippines, but this does not spread southwards into the main body of the archipelago[5]. The situation for dogs and chickens is if anything more perplexing; they do not seem to turn up in assemblages at all, until identified much later in Polynesia.

This is in marked contrast to the apparent evidence from linguistics. Blust (2002) conveniently summarises the linguistic evidence for faunal terms in Austronesian. He proposes:

Table 8: Domestic animal reconstructions in AN

species	level	proto-form
chicken	PMP	*manuk
cock	PMP	*laluŋ
dog	PAN	*asu/wasu
puppy	PAN	*titu
domestic pig	PAN	*beRek
? wild pig	PAN	*babuy

In the case of chickens and dogs this sharply contradicts the archaeological evidence; no chickens and dogs have yet been found at this period. Linguistically, they cannot be apical forms which gradually diversify through the Austronesian world, but represent either semantic shifts or widespread loanwords. If this is the case, how do we explain the reconstructions? In the case of chicken, **manuk* alternates with reduplicated forms which mean 'bird' and indeed in the putative branch of Austronesian represented by Tai-Kadai this is what it does mean (Benedict 1942; Ostapirat 2005). Blust assumes that 'bird' is the secondary meaning, but more likely this was the original meaning and it has shifted to chicken with the subsequent spread of the animal. The exact evidence for **laluŋ* is lacking but it occurs in the Northern Philippines and in the Lesser Sundas and is presumably either a widespread loanword or a semantic shift. The case of words for 'dog' is more perplexing, since this is well-attested in Taiwanese languages and widely in island Southeast Asia. However, surprisingly it has no reconstruction in proto-Oceanic. In addition, the same root occurs virtually all across Austroasiatic. Yet dogs are again conspicuous by their absence in the archaeology of island Southeast Asia. Dogs were probably domesticated from the Asiatic wolf and appear to be found in early sites in China (Savolainen et al. 2002). Could all those occurrences of the **asu* root represent a semantic shift? It seems unlikely. A useful clue is found in the fact that Tai-Kadai languages, despite their evident reflection of PMP in terms of basic numerals, do not have the **asu* term for dog. Indeed, it appears that the proto-form in Tai-Kadai (something like **hma*) is likely to be a borrowing from Hmong-Mien (see Table 10 below). If so, then this term may have been absent in PMP and all those occurrences of *#asu* are in fact loanwords, reflecting contact with Austroasiatic speakers (and the subsequent spread of the term once borrowed).

The case of the pig is still more perplexing. According to Blust (2002:93), reflexes of **beRek* occur in Puyuma, Tsouic, the Northern Philippines and some Borneo languages with the meaning 'domestic pig'. This is realised in Oceanic as **boRok*, a general term for 'pig'. As with 'dog' there is a widespread term in Austroasiatic, **C-liik* or **C-lek*, which is apparently cognate with Austronesian. The fricatives in Formosan languages may well be

5 There has apparently been an independent domestication of a highly local race on Lanyu (Orchid island) which may account for these finds (Keith Dobney, personal communication).

cognate with forms such as Pear *sru:k*. In this case, the Taiwanese domestic pigs probably came from the mainland of East Asia and the same source also donated the pig to Austroasiatic, hence these terms are related. However, Austronesian has another well-attested form for 'pig' **babuy*. This term often applies to wild pigs, but Blust (2002:93) finds the meaning of 'domestic pig' also widespread. As neither species is subsequently attested in the archaeological record, we have to assume the term devolved to the wild pig and was sporadically shifted back to the domestic pig.

Taro terminology is another complex story. There are two main complexes of terms for taro in Austronesian, **talef* and **ma*. In Southeast Asia there is a widespread term, *#traw?* which has reflexes throughout Austroasiatic (Table 5) but also appears to be cognate with Austronesian. Although Dempwolff (1938:128–9) reconstructed **talə(s)* for proto-Austronesian his evidence did not include either Taiwan or any languages near Formosa. This suggests that Austroasiatic speakers were the original domesticators of taro and that Austronesian speakers borrowed it during an early phase of contact, with the southern Philippines/Borneo the most likely zone for such contact. Since this cannot have been through contact with Negrito hunter-gatherers, Austroasiatic speakers may previously have been resident in insular Southeast Asia. Taro and other vegeculture had spread east from the mainland, and the expanding Austronesian speakers adopted it from the Austroasiatic speakers whom they subsequently assimilated, but not before borrowing their term for the plant. Ross et al. (2008:266) point out that reflexes of this root are rather scattered in Western Oceanic and that they are possibly borrowings from Eastern Oceanic, where the term is widespread.

The existing paradigm of Austronesian migration is crumbling in the face of a conspicuous absence of archaeological evidence for some of its central claims (for example Donohue and Denham 2010; Blench in press d). Its replacement will be a far more nuanced account of the movement of plants, animals and other types of material culture in the Southeast Asian region. The notion propounded by Bulbeck (2008) of 'fisher-foragers' and the emphasis placed by Solheim (1984/5) on trade may well be significant components in any new model. However, these views do not account for the extreme pervasiveness of the Austronesian languages, which must have replaced and assimilated a complex of different language families in numerous different sites in ISEA. This suggests that in addition to sailing technology and trade, the Austronesians must also have had a hugely attractive social, organisational and perhaps religious ideology which persuaded the residents of individual islands to adopt Austronesian culture. This would certainly explain the extraordinary diffusion of certain iconographic elements, such as the *bulul* figure, the *linglingo* and others, noted by art historians but not adopted by archaeologists. The expansion of Chamic on the Vietnamese mainland and the displacement or assimilation of Austroasiatic languages and cultures shows just how effective this type of cultural colonisation can be.

No consensus on a new paradigm for Austronesian is likely to be reached for some time. However, as with Sino-Tibetan, it is possible to outline a speculative model that at least can account for the interdisciplinary data. With all the usual caveats, the following hypotheses are put forward:

a) The Austronesian phylum must have a structure similar to that outlined by Blust, which has Taiwanese languages as an array of primary branches and Malayo-Polynesian constituting the remainder.

b) The series of nodes separating Oceanic from Proto-Malayo-Polynesian (PMP) that characterise earlier proposals must be dropped for lack of evidence. Western PMP looks like an array of primary branches, somewhat like Formosan.

c) The many Austronesian languages in islands with *in situ* residents must have spread largely by adoption rather than demographic spread.

d) However, this process was by wholesale language switch which would account for the limited evidence for non-Austronesian (NAN) substrates. In Melanesia, these processes broke down, hence both the phenotypic switch of Austronesians to Papuan physiognomy and the occurrence of a range of mixed languages. Traces of 'Papuan' structures can be detected in many Austronesian languages without this being evidence for prior settlement of uninhabited islands.

e) The primary means of cultural conversion was religious and socio-political, rather than through military conquest or commerce.

f) The material culture package supposed to be typical of Austronesians was in fact constructed from elements adapted along the way. Only when the Austronesians reach Polynesia and the Solomons do they propagate this package as an apparently coherent ensemble.

g) As an additional consequence, many of the reconstructions for subsistence-related terms in Austronesian hitherto considered solid must instead be mosaics of ancient loanwords, spreading either east from MSEA or west from Melanesia.

The Bellwood migrationist model has been enshrined in the prefaces of too many dissertations to be easily dislodged. But the mismatch with archaeology has now become too blatant too ignore. Restructuring Austronesian to take account of both its linguistic dominance and problematic presence in the archaeological record will become a major task for the immediate future.

7 Daic [=Tai-Kadai]

The Daic or Tai-Kadai languages, of which Thai is the most well-known and widespread representative, are spoken from southern Thailand into Laos, Cambodia, Vietnam and China. Up-to-date maps of their distribution are given in Edmondson and Solnit (1997a) who estimate the number of speakers of these languages as at least 80 million. Overviews of the phylum are given in Edmondson and Solnit (1988, 1997b; and Diller et al, 2008). The Daic languages represent a highly coherent grouping whose structure is well understood; a relatively long list of common glosses make it possible construct hypotheses concerning the subsistence and migrations of the proto-Daic speakers. All the most diverse Daic languages are in China: despite the marked southward extension of Thai today, the likely origin of Daic is in Guizhou (貴州). Despite this, the Daic expansion has no obvious archaeological correlate, although there are clearly cultural links with Austronesian speakers of Taiwan, for example dental ablation (Blench in press, b).

Surprisingly, however, there is no standard reconstruction of proto-Daic, although branches such as Kra, Tai and Hlai have lists of proto-forms (Li 1977; Hudak 2008; Ostapirat 2000; Norquest 2007). Figure 5 shows the internal relationships of Daic given by

Edmondson and Solnit (1997b) amended with reference to Thongkum (2001) and Ethnologue (2009).

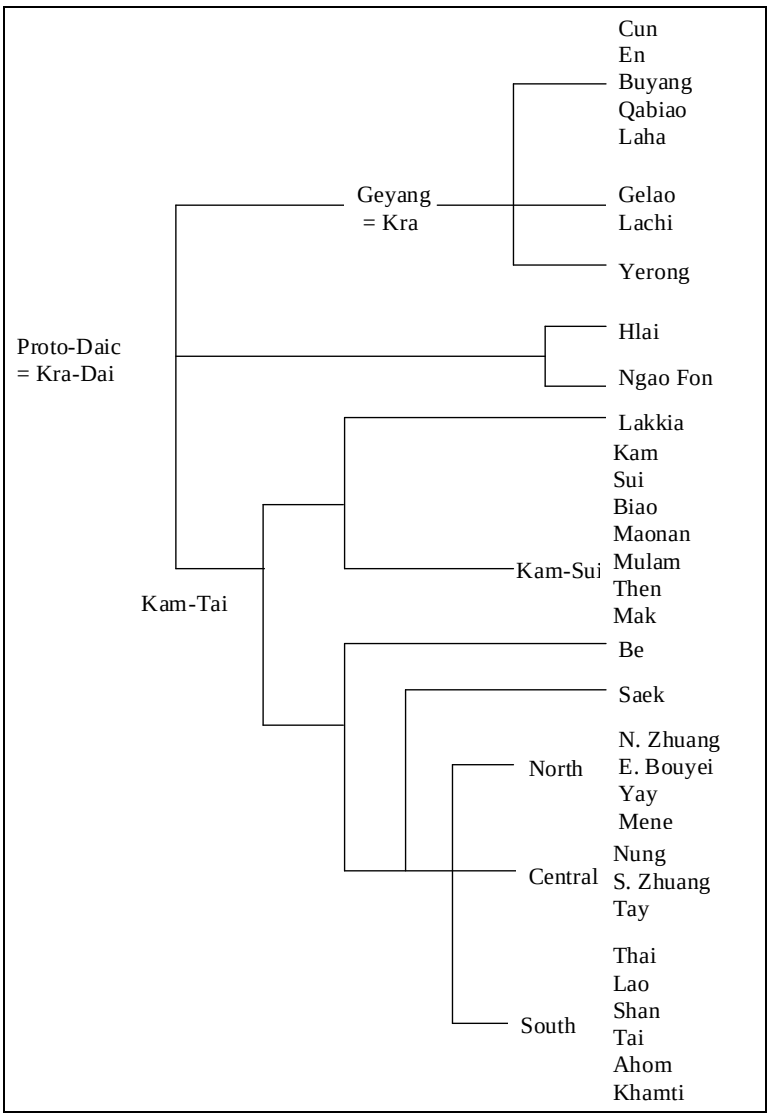

Figure 5: Daic subclassification.

The external affiliations of Daic have remained controversial, sharing as it does many features with surrounding language phyla, notably Austroasiatic, Hmong-Mien and Sino-Tibetan. These were used by Benedict (1942, 1975) to erect 'Austro-Tai', a macrophylum that would unite Austroasiatic, Hmong-Mien, Daic and Austronesian. Proposals for a linguistic connection between Tai and Austronesian date back at least to Schlegel (1901) and are extended in Wulff (1942). Baoya (1995) reviewed the rather extensive literature in Chinese (some of which argues for a genetic connection with Sinitic) and compiled a rather striking dataset comparing Daic with Malay. A failure to establish regular sound-correspondences meant that his observations have been passed over by later writers. Thurgood (1994) claimed that much of the evidence for hypotheses that link together the major language phyla of MSEA, such as Benedict's Austro-Tai, derive from ancient loanwords rather than genuine cognacy. However, Ostapirat (2005) set out a series of regular correspondences linking Daic with Austronesian, assuming a simple model of a primary split between Daic and Austronesian; in this model, the two would then be co-

ordinate branches. But this seems unlikely; Daic looks more like a branch of proto-Malayo-Polynesian and does not share in the phonological complexities of Formosan. Sagart (2004, 2005b) proposes that proto-Daic speakers migrated back to the mainland, to Guangdong and the region of Hainan island.

Norquest (2007:413) points out that the Hlai branch shares some striking lexical items with proto-Austronesian which do not occur in the other branches of Daic. A sample of the most convincing comparisons is shown in Table 9.

Table 9: Shared lexicon between Austronesian and proto-Hlai[#]

Gloss	Pre-Hl	PHl	PAn
slap	*pi:k	*phi:k	*pik
rub rope~weave	*bən	*pʰən	*bəl+bəl
pinch	*ti:p	*tʰi:p	*a-tip (PMP)
seven	*tu:	*tʰu:	*pitu
three	*ʈu:ʔ	*ʈʃʰu:ʔ	*təru
sharp	*ɟə:m	*tɕʰə:m	*ʈaɟəm
five	*ma:	*hma:	*rima
six	*nɔm	*hnom	*ʔənəm
maternal grandmother	*na:ʔ	*hna:ʔ	*ina ‘mother’s sister’
that	*C-na:	*C-na:	*i-naʔ
bury	*lɔmɦ	*hlomɦ	*ʈaləm
fish scale	*C-lə:p	*C-lə:p	*quʂəlap
eight	*ru:	*ru:	*waru
sell	*ri:wʔ	*ri:wʔ	*sariw

[#] Source: adapted from Norquest (2007:413)

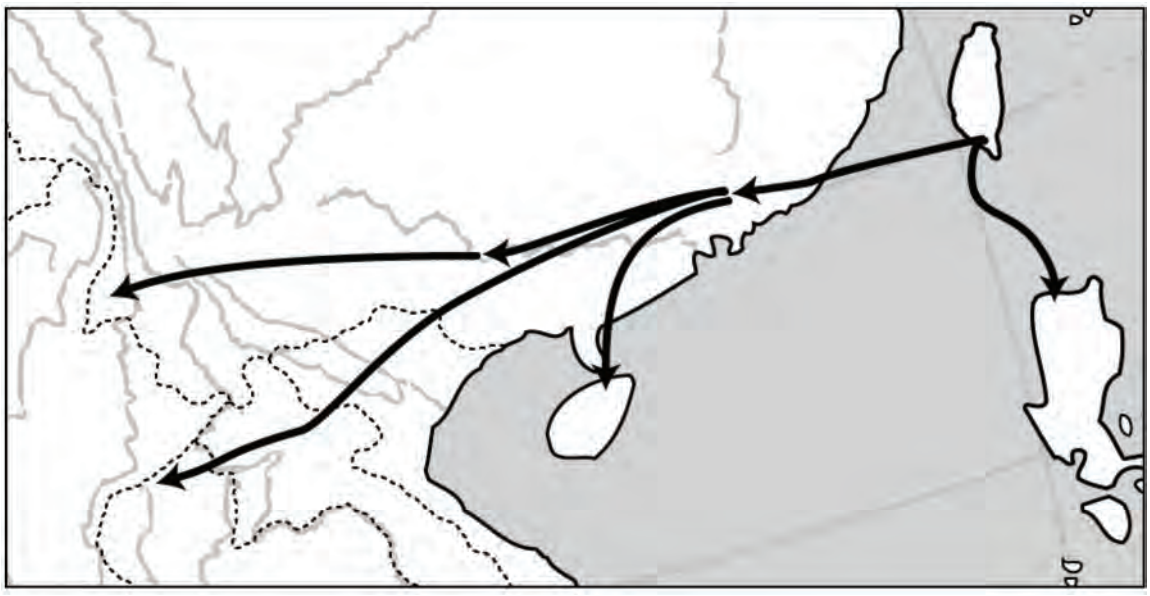

Figure 6: Hypothetical routes of Daic expansion.

This demonstrates neatly 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. The pattern of morphosyntactic reduction

is identical for the cognates with Kra pointed out by Ostapirat, namely the deletion or assimilation of the first syllable of the Austronesian form in Daic. None of these lexical items are specifically Formosan; they can just as well be PMP, which is certainly the case for Kra-Austronesian cognates identified by Ostapirat. The retention of these forms, in particular the numerals, is a striking testimony to the early diversification of Daic. Hlaic must have branched off at the same time as the Kra languages, retaining a specific set of Austronesian lexical items. Daic then became relexified and radically restructured following contact with Hmong-Mien and perhaps other languages of unknown affiliation. Such a migration would be around 4000 BP, in broad conformity with current dates for the first incursions in the Northern Philippines. Figure 6 shows the hypothetical routes of Daic expansion based on this evidence.

Daic languages are not all that diverse and 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' and 'dog' and at least some crops. Table 10 shows items extracted from Ostapirat relating to crops and domestic animals attested across all branches of Daic.

Table 10: 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)

Table 11: Proposals for the Daic subsistence lexicon

Item	Quasi-reconstruction	Possible source
taro	#pɣaak	< Taiwan names for *Alocasia macrorrhizos*
cooked rice	#mpVŋ	widespread mV- roots
husked rice	#saan	Sino-Tibetan and Austroasiatic
white rice	#rɔp	Daic innovation
millet	#pfeeŋ	Daic innovation
ginger	#kʰiŋ	< Proto-Hmong-Mien
buffalo	#kwaay	< Austroasiatic
goose	#ɣaan	< Austroasiatic

The exact referent of 'yam' is unclear; synchronically, this term can be applied to potato, sweet potato and taro. However, there is no doubt that the true yam, *Dioscorea* (probably *esculenta*), was formerly extensively grown throughout this region and has declined in recent times. Daic *#mpaw* for 'dog' is likely borrowed from Hmong-Mien **hmaŋ*^{C} 'wild dog' (Ratliff) as it resembles neither Austroasiatic nor Austronesian.

Table 11 presents some additional proposals for the Daic subsistence lexicon based on a more extensive compilation in Blench (in press b).

The argument for deriving the Daic name of taro from a Formosan term for *Alocasia macrorrhizos* is given in more detail in Blench (in press c).

The sheer variety of crops in this inventory strongly suggests that the proto-Daic speakers were established farmers. However, none of these terms except possibly sesame show links with Austronesian farming terminology. This was previously rather perplexing for the argument that Daic was a branch of Austronesian, but if it is the case that the Austronesians were marginal farmers emphasising fisheries and trade, the situation is more explicable. As Table 11 suggests, Daic does borrow terms from other regional phyla. Daic languages apparently underwent a lexical revolution with respect to agriculture on the mainland as a result of interaction with resident language phyla.

8 Language phyla and the antiquity of farming

Archaeological coverage of the region is highly skewed, and the quality and density of data from China has a tendency to bias interpretation. Nonetheless, it is apparent that farming began in the colder, drier zones north of MSEA proper. If the argument for the genesis of Sino-Tibetan is accepted, then its earliest phase was foragers in the foothills of the Himalayas, some exploiting sago and lowland fauna, others gathering wild cereals and montane animal species. An agricultural revolution took place ca. 6500 BP, stimulating a wave of expansion eastward into China. Agriculture then spread only slowly further south, presumably because tropical MSEA was such an abundant environment there may have been no need to farm.

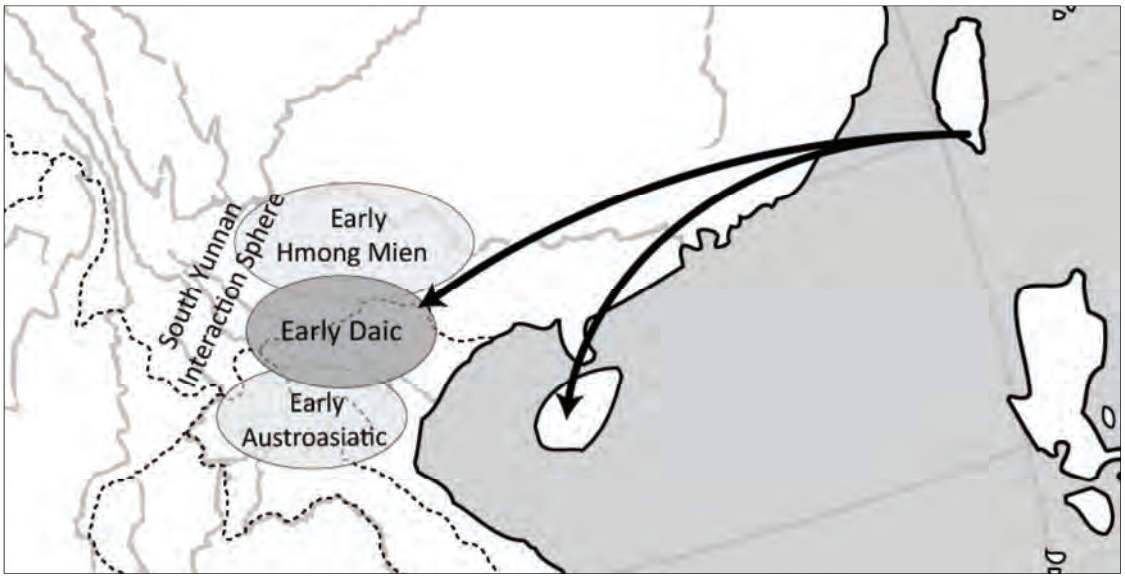

Figure 7: The South Yunnan Interaction Sphere.

However, by around 4300 BP there was a striking confluence of four quite distinct language phyla in what is now Yunnan and an almost simultaneous adoption of agriculture. This may well have been a result of the expansion of pre-Sinitic Sino-Tibetan speakers such as the Bai. Hmong-Mien, Austroasiatic and Daic speakers had all begun farming at the period when their respective proto-languages began to expand. This early period of intense interaction is provisionally named the South Yunnan Interaction Sphere

(SYIS) and a tentative graphic representation is shown in Figure 7. It cannot be emphasised too strongly that this remains a speculative hypothesis; only more detailed archaeology and linguistics will establish its credibility. It is suggested that the common roots in Sino-Tibetan reflecting agriculture also date from this era, although this must remain controversial due to a lack of properly presented data.

None of this implies that a Bellwood demic expansion model is necessarily appropriate, but it is improbable that these language phyla expanded significantly before the inception of agriculture. In other words, phyla cannot be significantly older than farming unless we reach for a model of 'extinct branches', subgroups in the foraging era which have now conveniently disappeared. People can have crops but expand for a variety of reasons; as is suggested in Sidwell and Blench (this volume), improved water transport stimulates dispersal. However, agriculture implies settlement and provide the preconditions for a variety of other important social changes, including technological advance, long-distance trade and more coherent military organisation.

There is a persuasive and pervasive stereotype that the languages of MSEA must be long established and that millennia of intensive interaction are responsible for the widely observed similarities between different language phyla as well as the remarkable interpenetration of fundamental vocabulary. However, a constellation of evidence from different disciplines suggests that this interpretation may be erroneous (Blench forthcoming). In fact it may be that the expansion of the major phyla is associated with the Neolithic and that the observed convergence can occur quite rapidly under specific conditions.

Much remains to be done to add weight to this revised scheme. We urgently await more credible reconstructions for Sino-Tibetan based on attestations from all the diverse branches. An evidence-based approach to the internal classification of Austroasiatic is essential. Direct material remains reflecting the transition to farming, based on actual archaeobotanical and archaeozoological materials would make claims for the dates of the inception of agriculture more plausible. Nonetheless, the potential to correlate different disciplines in creating an integrated prehistory of MSEA is clearly now within reach.

Acknowledgements

The origin of this chapter was first given as a paper at the Workshop 'Dynamics of Human Diversity in Mainland Southeast Asia' held at EFEO, Siem Reap, 6-10th January 2009 and I am grateful to the organisers, Nick Enfield and Joyce White, for the invitation to attend and present this. I benefited from discussions at the meeting but I would also like to thank Gérard Diffloth, Mark Post, Paul Sidwell and Laurent Sagart and the anonymous referees for additional comments. Thank you to Peter Johan Nijland for preparation of Figures 3, 6, and 7. Doug Cooper in Bangkok is doing remarkable work in making available source materials in electronic form and the data elements have benefited substantially from access to these databases. The present revision incorporates data from papers presented at subsequent conferences in 2009, including Austroasiatic (Bangkok), Sino-Tibetan (Chiang Mai) and the Indo-Pacific Prehistory Association (Hanoi). The argument concerning the origins of Sino-Tibetan was presented at the 16th Himalayan Languages Symposium, London September 2010 and the re-interpretation of Austronesian prehistory at the 13th EuraSEAA Meeting in Berlin in October 2010.

References

Abbi, Anvita. 2006. *Endangered Languages of the Andaman islands*. Munich: Lincom Europa.

Baoya, Chen. 1995. On the original relationship between Chinese and Kam-Tai. *Linguistics of the Tibeto-Burman Area*, 18(1). 149–171.

Bellwood, Peter. 1984/85. A hypothesis for Austronesian origins. *Asian Perspectives* 20. 107–117.

—— 1997. [2nd ed.] *Prehistory of the Indo-Malaysian archipelago*. Honolulu: University of Hawa'i Press.

—— 2005. *First Farmers: The Origins of Agricultural Societies*. Oxford: Blackwell.

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

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

—— 1972. *Sino-Tibetan: a conspectus ("STC")*. Contributing editor, James A. Matisoff. Princeton-Cambridge Series in Chinese Linguistics #2. New York: Cambridge University Press.

—— 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots.* New Haven: Human Relations Areas Files Press.

—— 1976. Sino-Tibetan: another look. *Journal of the American Oriental Society* 96(2). 167–197.

Blench, Roger M. 2005. From the mountains to the valleys: understanding ethnolinguistic geography in SE Asia. In Laurent Sagart, Roger M. Blench and Alicia Sanchez-Mazas (eds) *The peopling of East Asia*, 31–50. London: Routledge.

—— 2007. The language of the Shom Pen: a language isolate in the Nicobar Islands. *Mother Tongue XII*, 179–202.

—— 2008. Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? In Alicia Sanchez-Mazas, Roger Blench, Malcom D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 105–132. London: Routledge.

—— 2009. Remapping the Austronesian expansion. In Bethwyn Evans (ed.) *Discovering history through language. Papers in honour of Malcolm Ross*, 35-59. Canberra: Pacific Linguistics.

—— 2010. New evidence for the Austronesian impact on the East African coast. In Atholl Anderson (ed.) *Global origins and the development of seafaring*, 239-248. Cambridge: Macdonald Institute.

—— forthcoming. *Reconstructing the Southeast Asian past: archaeology, language, genetics and material culture*. Book ms. under consideration for publication.

—— in press a. Is Shom Pen a distinct branch of Austroasiatic? Paper given at the IVth ICAAL, Bangkok, November 2009 and in press for a special edition of Mon-Khmer Studies.

—— in press b. *The expansion of the Daic peoples: cultural and iconographic correlations.* Presentation at the XIVth EURASEAA meeting. Leiden, September 2008.

—— in press c. Vernacular names for taro in the Indo-Pacific region and their possible implications for centres of diversification. Paper for the special session on taro, IPPA, Hanoi. Special Issue of David Addison and Matthew Spriggs (eds) *Bulletin of the Indo-Pacific Prehistory Association*.

—— in press d. Austronesians or Austronesianisation? Paper presented at the 13th International Conference of the European Association of Southeast Asian Archaeologists, Berlin 27th September–1st October 2010 and now submitted for proceedings.

Blust, Robert A. 1976. Austronesian culture history: some linguistic inferences and their relations to the archaeological record. *World Archaeology* 8(1). 19–43.

—— 1995. The Prehistory of the Austronesian-speaking peoples: A view from language. *Journal of World History* 9(4). 453–510.

—— 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-Kuei Li (eds) *Selected papers from Eighth International Conference on Austronesian linguistics*, 31–94. Taipei: Academica Sinica.

—— 2002. The History of Faunal Terms in Austronesian Languages. *Oceanic linguistics* 41(1). 89–139.

Bodman, Nicholas Cleaveland. 1980. Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship. In Frans van Coetsem and Linda R. Waugh (eds) *Contributions to Historical Linguistics: Issues and Materials*, 34–199. Leiden: E.J. Brill.

Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.) *Tibeto-Burman languages of the* Himalayas, 1–72. Pacific Linguistics A-86. Canberra: Australian National University.

Bulbeck, David. 2008. An integrated perspective on the Austronesian diaspora: the switch from cereal agriculture to maritime foraging in the colonisation of island SE Asia. *Australian Archaeology* 67. 31–52.

Dempwolff, Otto. 1938. *Vergleichende Lautlehre des austronesischen Wortschatzes, vol. 3: Austronesisches Wörterverzeichnis*. Zeitschrift für Eingeborenen-Sprachen, Supplement 19. Berlin: Reimer.

Denham, Tim. 2004. The roots of agriculture and arboriculture in New Guinea: looking beyond Austronesian expansion, Neolithic packages and indigenous origins. *World Archaeology* 36. 610–20.

Diamond, Jared. 2001. Polynesian origins: Slow boat to Melanesia? *Nature* 410. 167.

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.

—— this volume, Austroasiatic word histories: boat, husked rice and taro.

Diller, Anthony V. N., Jerold A. Edmondson and Luo Yongxian (eds) 2008. *The Tai-Kadai Languages*. London: Routledge.

Dobney, Keith. Thomas Cucchi and Greger Larson. 2008. The pigs of Island Southeast Asia and the Pacific: New evidence for taxonomic status and human-mediated dispersal. *Asian Perspectives* 47. 59–74.

Donohue, Mark and Tim Denham. 2010. Farming and Language in Island Southeast Asia: Reframing Austronesian History. *Current Anthropology* 51(2). 223–256.

Dondrup, Rinchin. 1990. *Bugun language guide*. Itanagar: Arunachal Pradesh State Government Directorate of Research.

van Driem, George L. 1997. Sino-Bodic. *Bulletin of the School of Oriental and Southeast Asian Studies* 60(3). 455–488.

—— 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.

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

—— this volume, Rice and the Austroasiatic and Hmong-Mien homelands.

Dyen, Isidore. 1963. The position of the Malayo-Polynesian languages of Formosa. *Asian Perspectives* 7. 261–71.

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

—— (eds). 1988. *Comparative Kadai: linguistic studies beyond Tai*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

—— (eds). 1997a. *Comparative Kadai: the Tai branch*. 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: RoutledgeCurzon.

—— this volume, Linguistic diversity in mainland Southeast Asia.

Ethnologue. 2009. *A guide to the world's languages*. Dallas: SIL.

Ferlus, Michel. 1996. Du taro au riz en Asie du Sud-Est, petite histoire d'un glissement sémantique. *Mon-Khmer Studies* 25. 39–49.

Fuller, Dorian, Lin Qin and Emma Harvey. 2008. Evidence for the late onset of agriculture in the Lower Yangtze region and challenges for an archaeobotany of rice. In Alicia Sanchez-Mazas, Roger Blench, Malcom D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 40–83. London: Routledge.

Hajek, John. 1998. Kenaboi: an extinct unclassified language of the Malay Peninsula. *Mon-Khmer Studies* 28. 137–162.

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

—— 2002. *Early cultures of mainland Southeast Asia*. Bangkok: River Books.

—— 2004. Mainland Southeast Asia from the Neolithic to the Iron age. In Ian Glover and Peter Bellwood (eds) *Southeast Asia: from prehistory to history*, 41–67. Abingdon: RoutledgeCurzon.

Higham, Charles. F. W. and Thomas Higham. 2009. Dating Southeast Asian Prehistory: New Radiometric Evidence from Ban Non Wat. *Antiquity* 83. 125–144.

Hongo, Hitomi, Naotaka Ishiguro, Takuma Watanobe, Nobuo Shigehara, Tomoko Anezaki, Vu T. Long, Dang V. Binh, Nguyen T. Tien and Nguyen H. Nam. 2002. Variation in mitochondrial DNA of Vietnamese pigs: relationships with Asian domestic pigs and Ryukyu wild boars. *Zoological Science* 19. 1329–1335.

Hudak, Thomas J. 2008. *William J. Gedney's comparative Tai source book*. Honolulu: University of Hawai'I Press.

Lai Van Toi. 1999. The initial inhabitants of Co Loa [in Vietnamese]. *Khao Co Hoc* 3. 39–54.

Lewis, Helen, Victor Paz, Myra Lara, Huw Barton, Philip Piper, Janine Ochoa, Timothy Vitales, Jane A. Carlos, Tom Higham, Leee Neri, Vito Hernandez, Janelle Stevenson, Emil C. Robles, Andrea Ragragio, Rojo Padilla, Wilhelm Solheim II, and Wilfredo Ronquillo. 2008. Terminal Pleistocene to mid-Holocene occupation and an early cremation burial at Ille Cave, Palawan, Philippines. *Antiquity* 82. 318–335.

Li, Fang-Kuei. 1977. *A Handbook of Comparative Tai*. Oceanic Linguistics Special Publication #15. Honolulu: Univ. Press of Hawaii.

Matisoff, James A. 2003. *Handbook of proto-Tibeto-Burman*. Berkeley: University of California Press.

—— 2008. *The Tibeto-Burman Reproductive System: Toward an Etymological Thesaurus*. Berkeley: University of California Press.

Matsumura, Hirofumi, Marc F. Oxenham, Nguyen Kim Thuy, Nguyen Lan Cuong and Nguyen Kim Dung, this volume, Population history of Mainland Southeast Asia: The Two Layer model in the context of Northern Vietnam.

Meacham, William. 1984/1985. On the improbability of Austronesian origins in South China. *Asian perspectives* 26(1). 89–106.

Niederer, Barbara. 1998. *Les langues Hmong–Mjen (Miáo–Yáo)—Phonologie Historique*. Lincom Studies in Asian Linguistics 07. München: Lincom Europa.

—— 2004. Pa-hng and the classification of the Hmong-Mien languages. In Nicholas Tapp, Jean Michaud, Christian Culas and Gary Yai Lee (eds) *Hmong/Miao in Asia*, 129–146. Bangkok: Silkworm Books.

Norquest, Peter K. 2007. *A Phonological Reconstruction of Proto-Hlai*. Ph.D. Department Of Anthropology, University of Arizona.

Oppenheimer, Stephen. 2004. The 'Express Train from Taiwan to Polynesia': on the congruence of proxy lines of evidence. *World Archaeology* 36(4). 591–600.

Ostapirat, Weera. 2000. Proto-Kra. *Linguistics of the Tibeto-Burman Area* 23(1).

—— 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. 2010. *Hmong-Mien language history*. Pacific Linguistics. Canberra.

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*. 132–60. London: Routledge Curzon.

—— 1994a. Possible non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics* 33. 37–72.

—— 1994b. Unravelling the linguistic histories of Philippine Negritos. In Tom Dutton and Darrell T. Tryon (eds) *Language contact and change in the Austronesian world*, 443–475. Berlin: Mouton de Gruyter.

Rispoli, Fiorella. 2008. The incised and impressed pottery of Mainland Southeast Asia: following the paths of Neolithization. *East and West* 57. 235–304.

Ross, Malcolm, Andrew K. Pawley and Meredith Osmond. 2008. *The lexicon of Proto-Oceanic: The culture and environment of ancestral Oceanic society, Part 3: Plants*. Canberra: Pacific Linguistics, Australian National University.

Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444.

—— 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, Lingtuistics and Genetics*, 161–176. London: Routledge Curzon.

—— 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.

—— 2008. The expansion of Setaria farmers in East Asia: a linguistic and archaeological model. In Alicia Sanchez-Mazas, Roger M. Blench, Malcom D. Ross, Ilia Peiros and Marie Lin (eds) *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*, 133–158. London: Routledge.

Savolainen, Peter, Ya-ping Zhang, Jing Luo, Joakim Lundeberg and Thomas Leitner. 2002. Genetic Evidence for an East Asian Origin of Domestic Dogs. *Science* 298. 1610–1613.

Schlegel, Gustav. 1901. Review of Frankfurter's Siamese grammar. *T'oung Pao* 2. 76–87.

Shafer, Robert. 1947. Hruso. *Bulletin of the School of Oriental and African Studies* 12(1). 184–196.

Shafer, Robert. 1966/1967. *Introduction to Sino-Tibetan*. (Part I: 1966, Part II: 1967) Wiesbaden: Otto Harrassowitz.

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

Sidwell, Paul and Roger Blench, this volume, The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis.

Simon, Ivan M. 1970. *Aka language guide*. Itanagar: Government of Arunachal Pradesh.

Solheim, II Wilhelm G.1984/1985, The Nusantao hypothesis: the origin and spread of Austronesian speakers, *Asian perspectives* 26(1). 77–88.

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.

Starostin, Sergej A. 1991. On the hypothesis of a genetic connection between the Sino-Tibetan languages and the Yenisseian and North Caucasian languages. In Vitaly Sheveroshkin (ed.) *Dene-Sino-Caucasian*, 12–42. Bochum: Universitätsverlag Dr. N. Brockmeyer.

Szabó, Katherine and Sue O'Connor. 2004. Migration and complexity in Holocene Island southeast Asia. *World Archaeology* 36. 621–628.

Tayeng, Aduk S. 1990. *Sulung language guide*. Guwahati: Government of Arunachal Pradesh.

Terrell, John Edward. 2004. The "Sleeping Giant" Hypothesis and New Guinea's Place in the Prehistory of Greater Near Oceania. *World Archaeology* 36(4). 601–609.

Thongkum, Theraphan L. 2001. The Tai-Kadai peoples of Hainan island and their languages. In Kalaya M.R. Tingsabach and Arthur S. Abramso (eds) *Essays in Tai linguistics*, 189–203. Bangkok: Chulalongkorn University Press.

Thurgood, Graham. 1994. Tai-Kadai and Austronesian: the nature of the relationship. *Oceanic Linguistics* 33(2). 345–368.

Thurgood, Graham and Randy J. LaPolla (eds). 2003. *The Sino-Tibetan languages*. Routledge Language Family Series. London and New York: Routledge.

Vinh, Bui. 1991. The Da But culture in the Stone Age of Vietnam. *Bulletin of the Indo-Pacific Prehistory Association* 10. 1271–131.

Weidert, Alfons. 1987. *Tibeto-Burman Tonology: A comparative analysis*. Amsterdam and Philadelphia: John Benjamins.

White, Joyce C. 1995. Modeling the development of early rice agriculture: ethnoecological perspectives from northeast Thailand. *Asian perspectives* 34(1). 37–68.

White, Joyce C., Daniel Penny, Lisa Kealhofer and Bernard Maloney. 2004. Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. *Quaternary International* 113. 111–132.

Wolfenden, Stuart N. 1927. *Outlines of Tibeto-Burman Linguistic Morphology*. London: The Royal Asiatic Society.

Wulff, Kurt. 1942. Über das Verhältnis des malayo-polynesischen zum indochinesischen. *Det Künglige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 27(2).