

How many independent rice vocabularies in Asia ?¹

Laurent Sagart

CNRS, Paris, France

(Draft — not for citation)

1. The spread of rice as a multidisciplinary problem.

1.1. *The farming-language hypothesis, with a twist*

How can geneticists, anthropologists, archaeologists, cultural historians, linguists, arrive together at a deeper understanding of human prehistory in East and South Asia ? A hypothesis unifying our various problematics is that in that period, roughly the past 10 000 years, the expansions in East and South Asia of human populations, their languages, the cereals they cultivated, together with the genes of these cereals, all have the same underlying cause: the shift to agriculture and its demographic consequences. Populations of farmers can support larger families than hunter-gatherers, which gives them higher densities, and lets them expand with their genes, their crops and their languages. This is the well-known Bellwood-Renfrew farming/language hypothesis. This hypothesis is adopted here, with the standard caveat that not all linguistic expansions need to be agriculturally based (eskimo-aleut an obvious case); and with the twist that while agriculture *per se* will normally induce an increase in population density, making geographical expansion a potential outcome, it will not by itself suffice to lead to geographical expansion: another prerequisite is the possession of a diversified, versatile food procurement strategy, permitting adaptation to changing natural environments. As the environment changes, old techniques of food production may become inadapted and need to be replaced by new ones. Possession of a broad spectrum of food procurement strategies allows this to happen without an interruption in a group's food supply. In that sense, exclusive specialization in one kind of cereal is not conducive to wide geographical expansion.

1.2. *A multidisciplinary problem*

With rice, the multidisciplinary problem we face is to match each start of cultivation with a node dominating a clade in the genetic tree of varieties of *Oryza sativa*; with an archaeologically attested culture; and with a node in one of the region's language phylogenies. Unfortunately there is no

¹ The author thanks Bill Baxter, Michel Ferlus, Dorian Fuller, Susan McCouch, Martha Ratliff, Martine Robbeets, James Unger, George van Driem, Alexander Vovin, John Whitman and other colleagues for useful discussion and personal communications.

consensus on the number of starts of cultivation for rice. There are currently theories that say Asian rice, *Oryza sativa*, was put into cultivation only once (the "snowball model" of Vaughn et al.) and theories that say there were at least two (the Kovach-McCouch model, Kovach 2007): at least once for Indica-and-Aus, and at least once for Japonica-Javanica-Basmati. Londo et al. (2006) claim that indica was domesticated in south Asia and japonica in east Asia. Fuller argues that independent starts were made in two locations in the Yangzi Valley around 8.000 BP, and in the Ganges Valley maybe 6.000 BP (these are not the dates of domestication proper, which are later). There are still regions in southeast Asia (Burma, Laos, Cambodia) where yet undetected starts of cultivation may have taken place. The rice dendrogram in Garris et al. (2005) has 3 main clades: Indica, Japonica and Aus (Illustration 1).

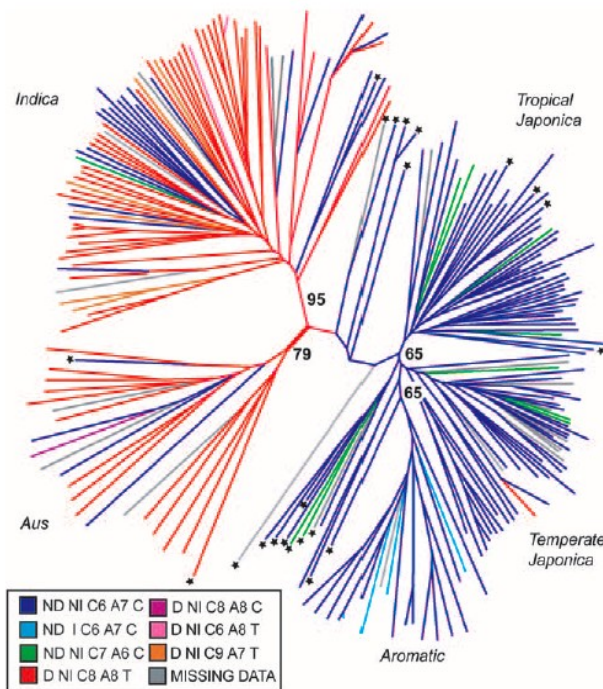


Illustration 1: Phylogeny of Oriza sativa according to Garris et al. (2005)

1.3. Starts of cultivation, seen by linguists

To linguists, a start of cultivation is a time during which a specific vocabulary is formed in a language to attend to the needs of speakers who cultivate and consume a particular plant. Some words may be inherited from the vocabulary of the ancestral hunter-gatherers who collected the plant in the wild: especially names for the parts of the cultivated plant —the stalk, leaves, ears, grains, husks, awns etc.. Others, such as as the name of the plant, will probably have to be created anew, since during a shift to cultivation gathering of the wild plant will continue for a while and

speakers will need two distinct names for the wild and the cultivated plant. For the purposes of this paper, the most interesting words are those whose earliest documented meaning is specific to rice.

1.4. Aims of this paper; transfers of vocabulary from earlier cultivated cereals; semantics of likely borrowings.

All the models we have for the expansion of rice suppose that rice cultivation spread across language boundaries. This paper aims, first, at documenting vocabulary transfers which can be taken as the linguistic side of transfers of rice cultivation; second, at identifying chains of language groups linked together by transfers of rice vocabulary; third, at identifying lack of rice vocabulary transfers between groups. The inference is that if languages show no evidence of having borrowed any rice-specific words from any other group, their speakers may have been involved in an independent start of cultivation.

Naturally, when rice cultivation spreads across a language boundary, we do not expect that the entire rice vocabulary of the donor language will be borrowed by the receiving language: new words can easily be created using the receiving language's own resources: for instance a new word for 'irrigated paddy field' may easily be created by compounding 'water' with 'field' or 'ground'. Especially if the speakers of the receiving language already cultivate another cereal, many words relating to that other cereal can be re-used for rice. But with certain notions a lexical borrowing is the more likely option. Especially with the name of the dehusked, ready-to-cook grain: a group adopting rice cultivation will have first become acquainted with ready-to-cook rice grains through trade, as an exotic food: objects of trade normally spread with their name.

2. East Asian language families with a reconstructible rice vocabulary

In certain East Asian language families, rice-specific terms may be reconstructed using the comparative method from the descendants back to the proto-language. The following language families and subgroups have a reconstructible vocabulary of rice-specific terms: Austronesian, Austroasiatic, Tai-Kadai (branch of Austronesian), Hmong-Mien (aka Miao-Yao), Sino-Tibetan, Korean, Japonic. The following do not: Tungusic, Mongolic, Turkic. This means that the speakers of proto-Austronesian (c. 3000 BCE), proto-Sino-Tibetan (c. 3500-4000 BCE), proto-Austroasiatic (date unknown, perhaps comparable to Sino-Tibetan), proto-Tai-Kadai (c. 1000-500 BCE), proto-Hmong-Mien (c. 500-200 BCE), Korean (?) and Japonic (c. 300-500 CE) knew rice and — presumably— cultivated it.

3. Vocabulary transfers among east Asian rice vocabularies

3.1. Japanese and Korean

The Japanese and Korean peoples cultivate temperate japonica varieties. The two languages may be genetically related (Whitman 1985; to appear), but while the inherited vocabulary that these languages share includes an agricultural component ('field', 'millet'), clear evidence for rice-related words is missing (Robbeets, p.c. July 2011). Japanese and Korean thus probably acquired rice cultivation after their separation (Unger 2008). Below we will see that the Japanese word *kome* 'dehusked rice' is a probable loanword from a pre-Austronesian language, suggesting that the linguistic ancestors of the Japanese acquired cultivation of (Japonica) rice from speakers of an eastern language within the macro-family I call Sino-Tibetan-Austronesian, with whom they were once in contact.

3.2. Austronesian

The Austronesian family is generally regarded as originating in a migration to Taiwan of fishing-and-farming groups from the mainland, c. 3500-3000 BCE. The proto-language first diversified in Taiwan; a migration out of Taiwan c. 2000 BCE resulted in the establishment, perhaps in the Philippines, of an Austronesian language ('PMP') ancestral to all the Austronesian languages outside of Taiwan. Knowledge of rice by the proto-Austronesians is widely recognized by linguists based on two reconstructed items: proto-Austronesian *Semay 'rice as food' and *beRas 'husked rice'. The latter includes a monosyllabic root² *-Ras with meaning 'fruit, flesh' etc. implying that at some point, perhaps before proto-Austronesian (but conceivably still in proto-Austronesian) the meaning was 'fruit, especially dehusked rice'. The Formosan vocabulary of millet has been under-recorded by investigators: it is possible that reflexes of *beRas mean 'millet grain' in more languages than is currently assumed.³ Large quantities of carbonized rice grains were discovered in 2002-2003 at Nan Kuan Li, a lowland site on the west coast of Taiwan dated c. 2800-2200 BCE (Tsang 2005), confirming linguistic reconstructions. The same site has also yielded carbonized grains of the millet *Setaria italica*, again in large quantities. A term for *Setaria italica*: *beCeŋ, had been reconstructed to proto-Austronesian. The Nan Kuan Li site attests to co-cultivation of rice and *Setaria italica* by the early Austronesians on Taiwan as early as the first half of the 3rd mill. BCE. Today western Austronesian peoples cultivate tropical japonicas and, in lowland locations, indica varieties. Both rice and millet were abandoned by the eastern (Oceanic) Austronesians as taro cultivation and

² Austronesian roots are meaning-associated syllables that recur at the end of independently reconstructible words, without the preceding syllable(s) being recognizable morphemes. They represent pre-proto-Austronesian monosyllabic words. See Wolff (2010?).

³ In Puyuma, a language of SE Taiwan, *bəras* means 'husked grain of rice or millet' (field notes, september 12, 2002).

arboriculture presented attractive alternatives. Reliance on fishing was never interrupted. Japonicas dominate among the traditional landraces maintained by the Austronesians in Taiwan; indica varieties seem to have been introduced into Taiwan after 1300 CE by Chinese settlers from Fujian, and even earlier in Borneo, before the Austronesian migration to Madagascar in the 1st millennium CE (below).

Visitors to Taiwan report no indigenous irrigated rice fields (outside of Chinese wet fields) until the Japanese occupation (1895-1945). Yet the Formosan vocabulary of rice shows that rice cultivation by the early Austronesians was not limited to upland dry fields: a proto-Austronesian root *-na 'flood-land' occurs in words meaning 'wet field' and 'riverside', suggesting lowland rice was cultivated on seasonal flood lands along rivers. The root occurs as a bound morpheme in Tsou *cxana* 'wet rice field' (analyzable as *cx*- < proto-Austronesian *CeNaq 'mud' plus *na 'floodland'), and as the second syllable in e.g. Paiwan *pana* 'river' (includes dry river bed, and low land along river) (Ferrell 1982); in Kavalan *Zena* 'field/wet field', etc. In addition there are indigenous words for 'rice seedling' and 'transplant rice seedlings' in the Tsouic languages (Tsuchida 1976:157), Bunun and Kavalan, although no proto-Austronesian term can be reconstructed.

3.3. Austronesian and Tai-Kadai

Based on shared innovations in the personal pronouns, numerals 5-10 and morphological innovations, Sagart (2004, 2005, 2009) argues that Tai-Kadai is a subgroup coordinate with PMP within the Austronesian family, which returned to the mainland after 2000 BCE. Since the Tai-Kadais are rice farmers, one expects that at least some of the rice vocabulary of Austronesian will be found in Tai-Kadai. Two items attest to this:

proto-Austronesian *-na 'floodland' ; Proto-Kra *na A 'rice-field' (Ostapirat 2000:229). Proto-Tai na: A 'paddy field' (Pittayaporn 2009).

Proto-Puluqish (a SE Formosan subgroup) *qaSaN 'rice in husks'. This is based on Amis *'asad* 'grains in husks mixed with rice' (Pourrias and Poinot, undated) and Paiwan *qasa:l*'chaff' (Ferrell 1982). Puyuma *asal* 'cut rice, before threshing' (Cauquelin 1991) must be a loan from either Paiwan or Amis (Expect Puyuma zero for proto-Austronesian *S). The final syllable in Papora *sesal*, *sisal* 'rice' (Ino 1998) also reflects proto-Austronesian -SaN, suggesting a 'root' *-SaN with rice-related meaning. To these, compare proto-Tai (Pittayaporn 2009) *sa:l A 'dehusked rice'. The sound correspondences between the proto-Tai and Austronesian forms agree with what is known

(Ostapirat 2005). The above etymology for proto-Tai *sa:l A is more likely, despite the difference in meaning, than the Austroasiatic origin suggested to the writer by M. Ferlus (p.c. 27/08/2002): the Tai word would originate in Khamou *hal* < earlier *sal 'to peel off', a form of limited distribution. Moreover the actions of peeling off (with a knife) and husking (by pounding) are very dissimilar from a motor point of view.

The Chinese word 秈 *sa[n] > sjen > xian1 used to mean 'indica rice' in modern standard Chinese matches proto-Tai *sa:l A perfectly and appears to be a Tai loanword into Chinese:⁴ there is no evidence that it designated indica rices, as opposed to tropical japonicas rices traded from Tai-Kadai speakers in south China. Large-scale contact between Chinese and early forms of Tai began c. 2200 years ago following the establishment in the region of present-day Guangzhou of the Chinese-led kingdom of Zhao Tuo.

After moving 'back' to the mainland, the Tai-Kadai came into intimate contact with Austroasiatic-speaking populations, borrowing part of their rice vocabulary (Ferlus, p.c. to LS, 27/08/2002):

- the general word for 'rice', proto-Tai (Pittayaporn) *C.qaw C is comparable with PMK (Ferlus) *rkoʔ / rŋkoʔ 'rice plant'.
- the word for 'swidden, dry field', proto-Tai *rɤj B (Pittayaporn) is comparable with Khamou hreʔ < *sreʔ, Khmer srae < srae 'dry rice field' as well as related Bahnaric words.

It is interesting that the Tai-Kadai name of the irrigated rice field is of Austronesian origin, while the name of the dry field is of Austroasiatic origin. This suggests that the Tai-Kadai moved to the mainland carrying lowland rice agriculture with them, and acquired upland rice cultivation from their new Austroasiatic neighbors. A problem is that Austronesian peoples in Formosa cultivate both highland and lowland rice. We propose the following explanation: the Austronesian expansions were led by fishermen who were looking for good fishing grounds in river estuaries, cultivating rice on the river sides by taking advantage of seasonal flooding, as a complement to fishing. These

⁴ The assertion is often made that the distinction between japonica and indica rices was known to the Chinese about 2000 years ago, as 粳 *jing1* (japonica) vs 秈 *xian1* (indica). While these meanings are those attached to these characters today, it is not at all certain that they designated japonica and indica varieties 2000 years ago. The word *xian1* first occurs in a now-lost version of the Fangyan, a c. 1 CE work on words occurring in languages of China other than standard Chinese, as quoted in the *Ji Yun*, an 11th century dictionary. It says <<江南呼粳為秈>> "Xian1 is the name of *Jing1* rice south of the Yangzi". As to 粳, it is defined under a slightly different graphic form in the *Shuo Wen*, a character dictionary of 120 CE as <<禾+亢, 稻屬。>> "Jing1 is a kind of rice". In a text from the Jin dynasty 265-419 CE we learn that *Jing* was dependent on irrigation: "Jing and *tú* rice are nourished by water and irrigation, while *Setaria* and *Panicum* are sown in upland fields" (晉左思<<魏都賦>>: "雨澍粳稌,陸蒔稷黍"). In a lexicographical text dated c. 543 CE we learn that *Jing* was non-sticky "jīng means non-glutinous rice 稻 (玉篇: <<粳, 不黏稻>>). It would appear, therefore, that the term 秈 *xian1* designated nonglutinous lowland rice from south of the Yangzi.

would be coastal people who did not practice cultivation of upland rice and did not have the corresponding landraces with them. The Austroasiatics, in contrast, were specialized in upland/swidden rice, and transmitted the technology, landraces and attendant vocabulary to the Tai-Kadai.

3.4. Sino-Tibetan

The Sino-Tibetan family is thought to be composed of two branches, Chinese vs. Tibeto-Burman (the rest), although competing proposals have been made in recent years (recently van Driem 1999, Blench and Post 2010). None is however supported by any linguistic innovations. A short list of potential innovations supporting the Chinese-vs.-Tibeto-Burman view of Sino-Tibetan phylogeny is given below. Old Chinese (first half of the 1st mill. CE) is the earliest reconstructible form of Chinese. Chinese rice is temperate japonica. Introduction of non-japonica varieties in early 11th-century CE China is historically documented (below).

Not all Tibeto-Burman groups cultivate rice. Especially on the Tibetan plateau and surrounding highland regions, buckwheat, domesticated locally, has supplanted rice. Buckwheat appears almost simultaneously in eastern Tibet c. 2600 BCE (Wang 1989) and in Xishanping in eastern Gansu c. 2600-2350 BCE (Li et al. 2007). Wheat, introduced from the west, is also found at Xishanping c. 2600 BCE (Li et al. 2007). The availability of these new grains made possible the abandonment of rice by some Tibeto-Burman groups. The following comparison, attested in two geographically and phylogenetically distant members of the family, implies knowledge of rice by the proto-Sino-Tibetans:

Old Chinese 𪎭 *C.m^s[e]j? > mejX > mǐ 'millet or rice grains, dehusked and polished'⁵, Proto-Bodo-Garo (Joseph and Burling 2006) *mai 1 'rice, paddy, cooked rice'.

The semantics point to rice grain in the final stages of processing, either ready for cooking or cooked. Chinese extended the term to millet grains, Bodo-Garo to rice in general.

There is concern, recently voiced by van Driem (2009), that the sound correspondences on Sino-Tibetan words for 'rice' may not be regular. Similarly Blench (2009) states that there is no evidence that the proto-Sino-Tibetans knew rice —indeed, that they were farmers. It is true that

⁵ This form is reconstructed with vowel /i/ in Baxter-Sagart v. 1.00; we now reconstruct *e because two words with 𪎭 as phonetic are read mjie and mjieX in MC; if it were *-ij? we would expect MC mjijX.

correspondences within Sino-Tibetan are not so well understood, in particular those relating to initial stop manner and to tone (Sagart 2006); in addition we do not have a reliable reconstruction of proto-Sino-Tibetan which would allow us to say that Old Chinese⁶ *C.m^s[e]jʔ* and proto-Bodo-Garo *mai 1 are the regular outcomes of proto-Sino-Tibetan *such-and-such. At least it can be said that good parallels can be found for all of the segmental and suprasegmental sound correspondences that this comparison implies—cognate decisions in Sino-Tibetan studies at this stage are based on nothing else—: for the initial, proto-Bodo-Garo *m- normally corresponds to Old Chinese *m- (Table 1); proto-Bodo-Garo main vowel *a sometimes corresponds to words with main vowel *e in Chinese (Table 2); Old Chinese final *-j normally corresponds to proto-Bodo-Garo *-i in closing diphthongs (Table 3); and proto-Bodo-Garo tone 1 and Old Chinese final *ʔ match in a significant number of forms (Table 4).

	proto-Bodo-Garo	Old Chinese
dream	jV3-maŋ	夢 *C.məŋ-s > mjuwngH > mèng ‘dream’
lose	Gu 1-ma	無 *ma > mju > wú ‘not have’
name	muŋ 1	名 *C.meŋ > mjieng > míng ‘name’

Table 1: Correspondence of Bodo-Garo initial *m and Old Chinese initial *m

	proto-Bodo-Garo	Old Chinese
change	*lai (no tone given)	易 *lek > yek > yì ‘change; exchange’
spirit, god	*mu-Dai 4	帝 *t ^h ek-s > tejH > dì ‘God’

Table 2: Correspondence of Bodo-Garo *ai with Old Chinese *e

	proto-Bodo-Garo	Old Chinese
die	thui 1	死 *sijʔ > sijX > sǐ ‘die (v.)’
buy	prai 1	買 *m ^s rajʔ > meaX > mǎi ‘buy’
break	phai 2	破 *p ^h aj-s > phaH > pò ‘break (v.)’

Table 3: Correspondence of Bodo-Garo *-i in closing diphthongs with Old Chinese *-j

⁶ The Old Chinese reconstructions used in this paper are based on the Baxter-Sagart system, ver. 1.00. Lists of reconstruction and explanatory files can be found online at: <http://crlao.ehess.fr/document.php?id=1217>.

	proto-Bodo-Garo	Old Chinese
die	thui 1	死 *sijʔ > sijX > sǐ 'die (v.)'
buy	prai 1	買 *m'rajʔ > meaX > mǎi 'buy'
ear	na 1	耳 *C.nəʔ > nyiX > ěr 'ear'
nine	k(h)u 1	九 *[k]uʔ > kjuwX > jiǔ 'nine'
rice beer	cu 1	酒 *tsuʔ > tsjuwX > jiǔ 'wine'
water	tui 1	水 *s.turʔ > sywijX > shuǐ 'water; river'

Table 4: Correspondence of Bodo-Garo tone 1 with Old Chinese *-ʔ

In view of the evidence in the above tables, it does not appear that the similarity of the Chinese and Bodo-Garo forms is the result of chance. The likelihood of it being the result of contact is not great either, in consideration of the geographical distance between these two.

Another Sino-Tibetan cognate set may be glossed as 'husked grain'. It cannot be taken as indicative of rice cultivation by the proto-Sino-Tibetans, although its Tibetan member has 'rice' as part of its meaning:

糲 *([m]ə-)rʰat > lat > li 'dehusked but not polished grain'

Written Tibetan ^mbras 'rice; fruit'

Dimasa *bere* 'to bear fruit'

The Tibeto-Burman part of the comparison is due to Benedict (1972), and to Matisoff (2003) for the Dimasa form. Dimasa, a close relative of Boro, has e for Tibeto-Burman *a in other forms, and regularly loses final *-s. The Chinese cognate was identified by Zhengzhang Shangfang. Note the nasal element in the onset in Chinese and Tibetan. It is unclear whether the -b- was lost in Chinese or gained in Tibeto-Burman. In either case, the nasal has to be old. Lushai raʔ < *ras 'fruit; to bear fruit' is clearly related but shows no trace of a labial pre-initial. Either the preinitial was lost or we are dealing with a related word at Sino-Tibetan level.

In addition, the Sino-Tibetan languages share a word for *Setaria italica*:

稷 *[ts]ək > tsik > jì 'millet (*Setaria italica*)'

Lhokpu *cək* '*Setaria italica*' (Gorge van Driem, p.c. to LS, June 25, 2004. Not phonologized),

Lepcha *č'ak* 'grain, food' (Mainwaring),

Written Tib. *č'ag* 'dry fodder for horses and other animals'⁷

The conjunction of rice and setaria allows one to locate proto-Sino-Tibetan relatively precisely in time and space: co-cultivation of rice and *Setaria* indicates a region situated between the Yangzi Valley, where rice was put into cultivation maybe c. 7000 BCE, and the Yellow River valley, where *Setaria* was domesticated perhaps c. 6500 BCE. As a result of parallel expansion, the two zones begin to overlap, and sites with the two cereals in domesticated form begin to appear in Henan in the second half of the 5th millennium BCE: Baligang in south Henan, c. 4200 BCE and Nanjiaokou in north Henan near the Yellow River, c. 3900 BCE: archaeologically this corresponds to the middle phase of the Yangshao culture. We equate 5th mill. BCE *Setaria*-and-rice Middle Yangshao with proto-Sino-Tibetan-Austronesian (below) and its *in situ* descendant proto-Sino-Tibetan with late Yangshao in the 4th millennium BCE. Breakup of proto-Sino-Tibetan into a Chinese and a Tibeto-Burman branch may be dated to the late 4th millennium BCE when the western Majiayao culture c. 3100-2700 BCE, in the upper Yellow river (Xishanping, c. 3000 BCE: rice, *Setaria*; later: wheat, buckwheat) separates from final Yangshao culture.

proto-Sino-Tibetan, like proto-Austronesian, has no specific word for 'rice field'. There is only one reconstructible item for 'field', something like *liŋ: Written Tibetan *zying* < *lying 'field, ground, soil, arable land' : 田 *l'iŋ > den > tián 'field; to hunt'. Outside of Written Tibetan the term also occurs in Lepcha *lyǎŋ*, Cuona *leŋ*¹³, Hayu *jing* 'dry field' (see the STEDT database), although some of these forms could be borrowed from Tibetan.

3.5. Objections to a N. China homeland for Sino-Tibetan

Identification of proto-Sino-Tibetan with a late stage of Yangshao is controversial. A tradition of research regards Sino-Tibetan as spoken by non-agricultural groups in central or southern Asia, and Chinese as intrusive in East Asia. Haudricourt and Strecker (1991) view the Chinese as originating in sheep herders from central Asia who became dominant over a Hmong-Mien speaking peasantry specializing in rice, borrowing their agricultural and commercial vocabulary. Sagart (1995b) showed that the alleged loans in the commercial vocabulary include characteristic Chinese morphology and must in fact be loans in the other direction. Sagart also rejected the views that the Chinese words for 'field' and 'flour' are borrowed from Hmong-Mien, but left open the possibility that the Chinese name of the rice plant is a very old borrowing from (an ancestor of) proto-Hmong-Mien into Chinese.

⁷ *Setaria italica* makes valuable fodder.

Starostin envisioned a Sino-Tibetan homeland in the Himalayan region; he assumed a Chinese migration east out of the Sino-Tibetan homeland after 3000 BCE. In his view (Starostin 2008) the archaeological Yangshao culture is best equated with proto-Altaic (Altaic is a much-disputed macro-family composed of Turkic, Mongolic, Tungusic, Japanese and Korean). However, the archaeologically documented presence in the middle Yangshao culture of rice and of a well-developed fishing component are not easily reconciled with the absence of corresponding terms in proto-Altaic: a more attractive match for proto-Altaic is the Hongshan culture in Liaoning province (Robbeets, p.c. to LS, July 2011), a culture of farmers of *Panicum* and *Setaria*, without rice, contemporary of Yangshao. Robbeets's views seem sound.

Starostin held that the pre-Chinese acquired agriculture at the end of their eastward migration: millet from the Altaic peoples, rice from the pre-Austronesians. In support of that view he listed some agricultural terms (2008) that he thought are not found in Tibeto-Burman, and that he treated as Altaic loanwords into Chinese. A prominent example is the name of the millet *Setaria italica*:⁸ 稷 *[ts]ək > tsik > jì. The source of the borrowing according to him is the PA reconstruction *ʒǰúgi 'Panicum miliaceum' in Starostin, Dybo and Mudrak (2003:1547-8). This reconstruction is based on proto-Turkic *yügür 'millet', proto-Tungusic *jiya / *jiye 'millet' and proto-Korean *cwok 'millet'. There are two serious problems here. First, this set does not match the sound correspondences observed on the more constrained set of Altaic etymologies discussed in Robbeets (2005).⁹ This means the set assembled by Starostin, Dybo and Mudrak is based on accidental resemblances; although the Korean form *cwok* 'Setaria' may be borrowed from Chinese. Second, the Chinese word *does* have good Tibeto-Burman cognates (above), showing knowledge of millet is Sino-Tibetan (Sagart 2008).

Opponents of a north China homeland generally argue that Sino-Tibetan phylogeny is poorly understood and that a primary split between Chinese and Tibeto-Burman (the rest) is not demonstrated; that the low linguistic diversity in north China is not what one expects of the homeland of an old and diversified family like Sino-Tibetan; Blench and Post (2010) further point out that some Sino-Tibetan groups in the Himalayan region do not rely on agriculture at all, which leads them to believe that the Sino-Tibetan languages acquired agriculture after the breakup of the

⁸ glossed by him as 'Panicum' apparently based on Li Huilin (1983:29) against the judgment of Bray (1984), Chang Kwang-chih (1980:147) and Fogg (1983).

⁹ Robbeets expects proto-Altaic *ʒ to be reflected as t- in Turkic, not y-; by her correspondences, no proto-Altaic vowel matches the array on the set by Starostin, Dybo and Mudrak; the medial consonant in Tungusic should be -g-, not -y-; and no explanation is provided for final -r in the Turkic forms.

proto-language. It is true that no detailed innovation-based model of Sino-Tibetan phylogeny has yet been presented. Yet for the basal structure of the family a Tibeto-Burman subgroup is supported by (1) the replacement of the proto-Sino-Tibetan word for 'dog', something like #kwial (= Old Chinese 犬 *k^hw^se[n]ʔ 'dog') by the earlier word for 'puppy', something like #kur/kuy (PTB *kwiy, proto-Austronesian *kurkur 'puppy'); (2) the partial or total merger of Sino-Tibetan *ə and *a into Tibeto-Burman *a. These points are disputed, however; (3) the pan-Tibeto-Burman loss of stop endings in words like 'sun': PTB (Benedict) *niy, Old Chinese *nit; 'varnish': PTB (Benedict) *tsiy, Old Chinese *ts^hit; 'arrow': PTB (Benedict) *b-la, Old Chinese 弋 *lək; 'change': PTB (Benedict) *lay, Old Chinese 易 *lek; etc.¹⁰ The low level of diversity in present-day north China is sufficiently explained by the levelling role of Chinese;¹¹ as to non-agricultural Sino-Tibetan-speaking groups in the Himalayan region, they may be groups who abandoned farming in favor of less labor-intensive subsistence strategies (as did Oceanic speakers in the Austronesian family); 'stranded' groups who somehow lost the technical know-how for agriculture (like the Tasaday in the Philippines, Reid 1993); or hunter-gatherers who shifted to the locally dominant language like Bantu-, Central Sudanic- and Ubangian-speaking Pygmies in Africa (Bahuchet 2006).

3.6. Sino-Tibetan-Austronesian

Sagart (most recent 2005) argues that the Austronesian and Sino-Tibetan families are genetically related as two branches of the Sino-Tibetan-Austronesian macro-family. This claim is predicated on the observation of sound correspondences on basic vocabulary and of morphological parallels with cognate markers. The sound correspondences obtain between the main syllable in Old Chinese words and the last syllable of Austronesian words. Sagart (2008) places the proto-Sino-Tibetan-Austronesian homeland in the same region as that for proto-Sino-Tibetan: this is based on the sharing by Sino-Tibetan and Austronesian of specific names for rice and millet, with the same correspondences as the rest of the shared vocabulary (Table 5, where Old Chinese is used as a proxy for Sino-Tibetan):

	Old Chinese	proto-Austronesian
rice, ready to cook/cooked	*C.m ^s [e]jʔ	Semay
<i>Setaria italica</i>	*tsək	beCeŋ

Table 5: two agricultural terms shared by Old Chinese and proto-Austronesian

¹⁰ For a longer list see Coblin (1986:31). Coblin reconstructs a glottal stop in proto-Sino-Tibetan, evolving to Old Chinese -t or -k depending on vowel context; but his proposal is not easily reconciled with the needs of tonogenesis in Chinese.

¹¹ Levelling of linguistic diversity by northern Chinese even applies to Chinese dialects, more diverse in SE China, despite the fact that the historical cradle of Chinese is in N. China.

The correspondence between proto-Austronesian *Semay and Old Chinese *C.m^f[e]j? are regular according to the sound correspondences in Sagart (2005):¹² all of m — m, a — e (also a — i) and y — j are well attested.

The correspondence between Old Chinese final -k and proto-Austronesian -ŋ in 'Setaria' above is unexplained: we would expect either k—k or ŋ — ŋ. Yet Old Chinese —proto-Austronesian parallels for k — ŋ exist. The same alternation is frequent within Sino-Tibetan and even within Chinese.¹³ Perhaps the Austronesian form incorporates a nasal suffix.

This must mean that proto-Sino-Tibetan-Austronesian was a precursor of proto-Sino-Tibetan in the same region, between the middle or lower Yangzi and the middle or lower Yellow river. We tentatively equate proto-Sino-Tibetan-Austronesian with Middle Yangshao, c. 4500 BCE —the date of the earliest layer in Baligang, assuming that other similar sites will see the light in the same region or closer to the eastern seaboard at similarly early, or even earlier dates. Rice remains from earlier sites like Jiahu in the same region are either from rice collected in the wild or not fully domesticated (Fuller et al. 2007).

Rice cultivated by proto-Sino-Tibetan-Austronesian speakers must have been japonica, since both the Sino-Tibetans and Austronesians are japonica farmers. Separation of tropical and temperate japonicas probably occurred after the separation of the two branches of proto-Sino-Tibetan-Austronesian. The Sino-Tibetan-Austronesian theory gives a natural explanation to the sharing of japonica rice by the Sino-Tibetans and Austronesians.

Since Setaria is sacred both among the early Chinese and the Austronesians in Taiwan but rice is not, it makes sense to suppose that the ancestors of proto-Sino-Tibetan-Austronesian speakers were Setaria farmers and pig raisers without cultivated rice, who had already developed religious practices centered around Setaria when they acquired rice. I tentatively equate this millet-only ancestor of proto-Sino-Tibetan-Austronesian with early Yangshao, and beyond, with the Cishan-Peiligang culture. Archeologically the appearance at Karuo in Eastern Tibet of grains of Setaria italica with a calibrated C-14 date range of 5555-4750 BP (Fu Daxiong 2001:66) possibly represents an early westward expansion of the first Setaria-based cultures in North China, an

¹² Even though Sagart (2005) gives the Old Chinese vowel as *i.

¹³ This kind of alternation is known as 陽入對轉 to practitioners of traditional Chinese phonology and etymology.

expansion which left no linguistic traces: the date is too early to associate with Tibeto-Burman.

Because co-cultivation of *Setaria* and japonica rice is characteristic of the Sino-Tibetan-Austronesian expansion, one may suppose that acquisition of japonica rice as a second cereal with different humidity requirements from *Setaria* was crucial in accelerating demographic growth and geographical expansion; domesticated pigs (probably fed *setaria*-derived products), fishing, hunting and gathering provided the necessary adaptability component.

It is noteworthy that proto-Austronesian *beRas 'dehusked rice', Old Chinese 糲 *([m]ə-)r^sat > lat > li 'dehusked but not polished grain' and Tibeto-Burman words like Written Tibetan mbras 'rice; fruit' correspond phonologically according to the correspondences given in Sagart (2005). This means that the word was part of the proto-Sino-Tibetan-Austronesian language, although it may have meant no more than 'fruit; dehusked grain of cereal', as Tibeto-Burman and Austronesian independently attest.

Neither proto-Austronesian nor Sino-Tibetan have a specific word for 'rice field' although both have a generic term: the same is implied of proto-Sino-Tibetan-Austronesian. We have seen that the early Austronesians may have cultivated lowland rice on floodable river sides, which we have suggested were called *na; this is, as we saw, the inherited Tai-Kadai word for the wet rice field. It is noteworthy that a term matching this *na in pronunciation and meaning occurs in Sino-Tibetan: Old Chinese 洳 *na-s > nyoH > rù 'wet ground', Written Tibetan na 'meadow'. It is possible that rice was first cultivated on floodland along riversides in north China, before the technology of irrigation was adopted.

A model of the expansion of the eastern component of the Sino-Tibetan-Austronesian macro-family was presented in Sagart (2008), in which much earlier dates for proto-Sino-Tibetan-Austronesian were given. The dates given here take Fuller's (2008 ?) revision of the chronology of rice domestication into account. Under Fuller's younger chronology, loss of shattering is first attained in two loci in the Yangzi valley by 4500 BCE. Accordingly the present model has the original Sino-Tibetan-Austronesian/middle Yangshao nucleus with domesticated rice and millet expanding east, reaching the eastern seaboard in south Shandong, there evolving into the *Setaria*- and rice-based Dawenkou culture c. 4000- BCE. From there, seaborne groups expand rapidly along the coast in a southerly direction, reaching Fujian in the 4th millennium, and crossing to Taiwan c. 3500-3000 BCE. In this model, Dawenkou culture farmers spoke a sister language of proto-Sino-Tibetan,

ancestral to proto-Austronesian, which would have had for 'rice ready to cook/cooked' a cognate of proto-Austronesian *Semay, a point further discussed in the next section.

3.7. *Transmission of rice by the pre-Austronesians to the pre-Japonics ?*

There is growing agreement among Japanologists that the settlement of Japan by Japonic speakers was effected in the second half of the first millennium BCE by a people ('Yayoi') from the Korean peninsula, where evidence for a now-extinct language ('Old Koguryo') related to Japanese can be detected in toponyms. Beyond Korea, Unger (2008) places the pre-Japonics in Shandong. Robbeets (p.c., July 2011) also has in mind a Shandong origin of the pre-Japanese. My own placement of the pre-Austronesians in south Shandong's Dawenkou culture (Sagart 1995a) makes them close neighbors of the pre-Japonics. While arguments for a genetic relationship between Japanese and Austronesian are not credible (Vovin 1994), contact between them is a possibility. The practice of tooth evulsion (extraction of the lateral incisors in boys and girls as a puberty rite) supports this: it originated in Dawenkou c. 4000 BCE (Han and Nakahashi 1996) and is found in exactly the same form among the early Formosans, e.g. in all adults skulls at the Nan Kuan Li East site in western Taiwan, c. 2800-2500 BCE (Pietruszewsky et al. 2009).¹⁴ Brace and Nagai (1982) describe tooth evulsion among the Yayoi people in Japan and propose that the custom was passed on to contemporary Jomon people, who (they argue) regarded that custom as a manifestation of higher civilization. If the pre-Austronesians and pre-Japonics were in contact in Shandong, this furnishes an opportunity for the transmission of both rice cultivation and tooth evulsion from the former to the latter. A likely linguistic signature of the cultural transmission of rice is the Japanese word *kome* 'dehusked rice', Old Japanese *kome₂, proto-Japonic *kəmai or *kəmāi. Neither of these proto-Japonic forms look like native Japanese words, the first because it would violate Arisaka's laws on cooccurrence of vowels, and the second because one would expect to find komo- alternating with *kome* in modern compound words (I am grateful to M. Robbeets and to J. Unger for these explanations). Proto-Japonic had no h- sound and treats foreign /h/ as k: therefore a possible source of proto-Japonic *kəmai or *kəmāi is a foreign *həmai or *həmāi. This is very close to proto-Austronesian Semay, if one assumes that the sibilant at the beginning of this word changed to h-, a frequent change cross-linguistically.¹⁵

¹⁴ "Tooth ablation, in this case most likely a rite of puberty, was observed in all of the adult Nankuanli East individuals examined in this study. With a single exception, the pattern observed was the intentional removal of both maxillary lateral incisors and canines well before the time of death." (slide 19)

¹⁵ Amis (SE Taiwan) *hmay* 'cooked rice' is treated by Blust and Wolff as the outcome of proto-Austronesian *Semay. However proto-Austronesian *S should give s- in Amis; Amis h- reflects proto-Austronesian *h. The Amis word could go back to a proto-Austronesian *hemay, identical with a possible source of J. *kome*.

3.8. *From whom did the proto-Sino-Tibetan-Austronesians get rice ?*

Judging from their location in the mid-Yangzi Valley at the time they enter history, the Hmong-Miens are prime candidates for the role of descendants of the first domesticators and donors of rice to their northern neighbors. This location suggests a filiation with the Qujialing culture, located in Hubei, Hunan and parts of Jiangxi c. 3000-2600 BCE. Qujialing, like its possible predecessor Chengtoushan (4500-3000 BCE), had a rice-based economy (although *Setaria*, probably acquired from the north, was also found at Chengtoushan). The north Hunan culture typified by Chengtoushan in its early stages is old enough to be the source of the transmission of rice north of the Yangzi, into proto-Sino-Tibetan-Austronesian. Yet, in marked contrast to Austroasiatic, original elements in the rice-specific vocabulary of the family are surprisingly few. Probably indigenous is the Hmongic term proto-Hmongic *ntsuw C 'dehusked rice' (Ratliff 2010:242). Most of the other rice-specific terms that can be reconstructed to Hmong-Mien are either clear loans from Chinese or have been suspected of being borrowed from Chinese.

A point to keep in mind is that, unlike Sino-Tibetan, Austronesian or Austroasiatic, the Hmong-Mien family is very young and that the reconstructed proto-Hmong-Mien lexicon includes a large number of Chinese loanwords which have the same sound correspondences across Hmong-Mien languages as the specifically Hmong-Mien vocabulary. These words, which were evidently part of the Hmong-Mien proto-language, show phonological features which are characteristic of late Old Chinese, roughly 500-220 BCE¹⁶, a period of strong cultural dominance of Chinese in the region. Some of these words relate to technical innovations of the same period.¹⁷ Even rice-related terms were borrowed from Chinese (see next section for examples). Four millennia, therefore, separate proto-Hmong-Mien from the domestication of rice: the proto-Hmong-Mien vocabulary of rice can only be a very indirect reflection of the original Yangzi rice vocabulary.

3.9. *Hmong-Mien rice vocabulary borrowed into Chinese ?*

Haudricourt and Strecker (1991), Ratliff (2010), van Driem (2009) among others have suggested that some Hmong-Mien rice vocabulary items were borrowed by Chinese specifically.

Haudricourt and Strecker (1991) claimed a Hmong-Mien origin for 秧 'young rice plant, seedling'; 稻 'rice plant'; 粉 'flour'; 糍 'bread, pastry'; and 買 'buy' and 賣 'sell'. Sagart (1995) showed the

¹⁶ Sound correspondences relating Old Chinese with the earliest layer of Chinese loans to Hmong-Mien can be found at <http://crlao.ehess.fr/document.php?id=1217>.

¹⁷ So strong was Chinese pressure on Hmong-Mien in the late Old Chinese period that 35% of the reconstructed Hmong-Mien lexicon in Wang and Mao (1995) is of Chinese origin (Ratliff 2010:227, n. 87).

direction of borrowing to have been unambiguously from Chinese into Hmong-Mien in the case of 'buy' and 'sell', and gave arguments that the word for 'flour' was also loaned by Chinese to Hmong-Mien. The term 秧 *yāng*, MC 'jang (= ʔjaŋ) 'young rice plant, seedling' is attested textually very late in Chinese (Tang dynasty, 618-907 CE), yet does have a clear Chinese-internal etymology: it is an infix-less cognate of 英 *ʔ<r>aŋ > 'jaeng > yīng 'young grass plants at the growing/flowering stage, before producing seeds'.¹⁸ In Hmong-Mien, in contrast, Ratliff (2010) reconstructs two phonetically similar but historically non-cognate forms in proto-Hmongic and proto-Mienic: *ʔjeŋ A and *ʔjaŋ A respectively: this makes it likely that Hmongic and Mienic borrowed the term separately, in the 1st millennium CE, from Chinese. With 稻 *[[ʔ]uʔ, the word designating the rice plant in modern Chinese, the oldest tokens of the corresponding character (Zhou bronze inscriptions, 1st mill. BCE) are written with the signific 米 'dehusked grains', instead of 禾 'cereal plant' in the modern graph. The word appears therefore to have originally referred to rice grains before shifting its meaning to 'rice plant'. Etymologically it is based on a verb 𪗇 *[[ʔ]u, *[[ʔ]uʔ, [[ʔ]uʔ (several readings) 'to scoop out hulled grain from a mortar', which also forms the right part of the character 稻. This confirms the evidence that the meaning 'rice plant' is secondary and very probably identifies the noun's earliest meaning as 'rice grains out of the mortar, after dehusking'. The square brackets around the initial consonant in the reconstructed Old Chinese form are there to indicate that several Old Chinese sources would give the same MC outcome as Old Chinese *lʰ. One of these is Old Chinese *m.lʰ, which means that one possible Old Chinese reconstruction of 稻 is *m.lʰuʔ. This may support the idea of a connexion with the proto-Hmong-Mien word *mbləu A 'rice plant'. If so, the fact that the Hmong-Mien form has the derived Chinese meaning argues for a loan from Chinese into Hmong-Mien rather than the reverse. The correspondence of rhymes, proto-Hmong-Mien *-əu — Old Chinese *u is compatible with a loan from Chinese: Ratliff (2010:119) cites the parallel of 媪 *ʔəuʔ 'old woman' corresponding to PHM *ʔəuX 'elder sister/wife'. However, in 'rice' one would expect tone B in proto-Hmong-Mien form opposite Old Chinese final -ʔ. Tone A in the proto-Hmong-Mien form requires a Chinese variant without final -ʔ. This is not impossible since the verb on which the noun is based does have such a variant, cf. above. On balance, the likelihood is that the word is a Chinese loan into Hmong-Mien. At any rate the semantics clearly show that the word is not a Hmong-Mien loan into Chinese.

Ratliff (2010:243) cites the proto-Hmong-Mien word *hnɔn 'grain head' which has an extended

¹⁸ The Shuo Wen, a Chinese character dictionary of c. 100 CE, gives the following gloss for 英: <<草榮而不實者>> "grasses at their developing/blooming stage, before producing grains".

meaning as 'bag/pocket'. She compares this with 囊 *n^haŋ > nang > náng 'sack, bag', proposing this Chinese term is a loan from Hmong-Mien —since it occurs in Chinese only in the derived Hmong-Mien meaning—. Alternatively, the two words could be look-alikes, as implied by the anomalous correspondence between proto-Hmong-Mien *-n and Old Chinese *-ŋ.

Van Driem (2009) argues all of the Old Chinese words 秠 *mə.lut 'glutinous millet', 田 *l^hiŋ 'field', 鎌 *[r]em 'sickle', 糍 *[g](r)a(k)-s 'cakes', 擣 *t^huʔ 'pound, thresh' and 甑 *s-təŋ-s 'steamer' are loans from Hmong-Mien. This cannot be true of 'sickle': if proto-Hmong-Mien *l^him 'sickle' had been borrowed by Chinese before the 4th century BCE, when it first occurs in the *Mozi*, a philosophical work, it would have been subject to the change of Old Chinese *l into y- in the 1st century CE and would appear in Middle Chinese with initial y-, not l-. It is also demonstrably false for 'steamer', 甑 *s-təŋ-s, an instrumental noun with s- prefix of the verb 蒸 *təŋ > tsiŋ > zhēng 'to steam (v.t.)'. The onset in 'steamer' changed from *s-t- to ts- between Old Chinese and MC and the Hmong-Mien form was evidently borrowed after that change. The word 秠 *mə.lut 'glutinous millet' is part of a correspondence set where Hmong-Mien *mbl- corresponds to Old Chinese *mə.l-, also including 'tongue' and 'eat'; these words in all likelihood have the same history of borrowing: but it is implausible, given the cultural, economical and military pressure of Chinese on its southern neighbors in historical times, that Chinese borrowed 'tongue' and 'eat' from Hmong-Mien, although the reverse is possible. Borrowing of the word 'eat' from Hmong-Mien into Chinese can further be excluded on semantic grounds: the Hmong-Mien word proto-Hmongic *mbljæ C 'have food with rice' specifically excludes rice and surely belongs to the Chinese higher cultural component. The source of the borrowing is 食 *mə-lək > zyik > shí 'eat', the standard Chinese word in this meaning, a word moreover with a good Sino-Tibetan etymology: Bradley (1979:#630) reconstructs proto-Loloish *m-lyak L 'to lick'. The word 'pound' is onomatopoeic: the resemblance between the Hmong-Mien and Chinese forms is not necessarily explained in terms of inheritance or contact. Phonologically 'field' and 'cakes' could be loans in either direction, but (like the other forms in van Driem's list) they are not rice-specific terms, which weakens the argument that they are loans from Hmong-Mien. 'Field' moreover has Tibeto-Burman cognates (above).

3.10. *Austroasiatic*

The Austroasiatic family is composed of languages spoken between Vietnam and the Indian subcontinent. Its geographical unity has been dislocated by the expansions of Indo-Aryan, Tibeto-Burman, Tai-Kadai and Austronesian languages: this implies that the Austroasiatic family was

already geographically spread out while these expansions were underway. Unlike Hmong-Mien, Austroasiatic was shielded from Chinese influence by its southerly location. The age of the family is not known but the impression is one of substantial time depth, perhaps broadly similar to that of the Sino-Tibetan family. The traditional view of the family's structure opposes a western group: the Munda languages, spoken in eastern and central India, to the rest ('Mon-Khmer'). This view is increasingly called into question as a convincing body of uniquely shared Mon-Khmer innovations has never been presented (Sidwell 2009). The location of the Austroasiatic homeland is much discussed. It is broadly agreed that Munda linguistic typology shifted from a southeast Asian type to a south Asian type. This suggests an adaptive change following a migration from southeast to south Asia. Stampe and Donegan (REF) think language contact not necessary to explain the shift, arguing plausibly that the change in overall typology was triggered by a single shift in speech rhythm: this leads them to doubt that the Austroasiatic homeland was in southeast Asia. Even if they are right, language contact would still be useful in providing a motivation for the initial shift in speech rhythm. Sagart (to appear, 2001) links Anderson's observation (2004) of a particular pattern in Austroasiatic languages including Munda whereby monosyllables have their main vowel doubled and the two vowel parts have a glottal stop inserted between them to obey a constraint that words have to be bimoraic, to the same strategy in Austronesian languages of Taiwan and the Philippines. Since this shared pattern is unrelated to rhythm, it reinforces the view of Austroasiatic as an east Asian language group. Diffloth (2005) argues for a tropical homeland, perhaps near the bay of Bengal, on the ground of his reconstruction of names of tropical fauna to proto-Austroasiatic. The supporting evidence has not been published. If Munda turns out not to be a primary branch of Austroasiatic, this will shift the center of gravity of the family, and its homeland, further towards the east.

Diffloth (2005) gives a list of proto-Austroasiatic rice-specific terms (Table 6), the supporting evidence for which has not yet been presented:

rice plant	#(kə)ba:ʔ
rice grain	#rəŋko:ʔ
rice outer husk	#cəŋka:m
rice inner husk	#kəndək
rice bran	#phe:ʔ
dibbling stick	#jərmuəl

Table 6: proto-Austroasiatic rice-specific vocabulary (Diffloth 2005)

'Rice bran' #phe:ʔ has a certain resemblance with proto-Hmong-Mien *mph̥iək 'chaff/husk' but it is not clear whether this should be regarded as a chance resemblance or a meaningful one. Other than that, the only significant match between this set and another language group is the term 'rice grain': #rəŋko:ʔ which, as we have seen, was borrowed by Tai-Kadai.

Diffloth notes the absence of terms relating to irrigated rice cultivation such as the wet rice field. An interesting item in this connexion is 'dibbling stick', a tool used by the more conservative MK groups to make holes in the ground in which rice seeds are planted. Unless this term is not rice-specific, it appears to point to rice cultivation in dry fields.

Ferlus (2010) lists other, more geographically restricted forms: proto-Katuic *s-rɔ: 'paddy, raw rice', also found in Mon, Khmer and Sora (Munda) and a form *s-ŋɔʔ 'paddy, raw rice' found in Palaung-wa, Khmu and Mon. In addition he identifies minor forms such as *cəh, *ha:l and *sa:, all 'raw rice, paddy'. None of these appears to have been borrowed from an outside source, or to have been loaned to an outside group.

Knowledge of rice by the PAustroasiatics naturally raises the question whether rice cultivation in south Asia ultimately goes back to a start of rice cultivation by Austroasiatic-speaking peoples. This could have happened in either of two ways: under the hypothesis of a south Asian homeland, the PAustroasiatics could be the first domesticators of rice in north India. Under the hypothesis of a southeast Asia homeland, the Mundas could have migrated west, carrying with them rice cultivation.

The first theory is illustrated by Kuiper's and Witzel's work. Kuiper (1948, 1950) identifies a Munda substratum in the Rgveda, based on what he interprets as prefixation patterns. Witzel (2000) follows Kuiper, speaking of "para-Munda", meaning a now-extinct western branch of Austroasiatic. Krishnamurti (2003:38) comments that "the main flaw in Witzel's argument is his inability to show a large number of complete, unanalyzed words from Munda borrowed into the first phase of the Rgveda". In the absence of specific Austroasiatic words, identification of a Munda-related language as the prefixing substratum language in the Rgveda is doubtful. Other prefixing candidates are Tibeto-Burman and Burushaski. A Tibeto-Burman presence in NW India at the time of the Rgveda should not at all be deemed impossible; Burushaski probably belonged to a significantly diversified language family before becoming an isolate; a third possibility is an extinct prefixing language. In support of a Munda role in south Asian agriculture, Witzel (1999) argued that the proto-Koraput Munda form *ə-rig 'Panicum miliare' (Zide and Zide 1976) is the source of the main south Asian terms for 'rice': Ved. *vr̥hi* and of apparently related terms for rice in Iranian and Dravidian languages: Persian *birinj*, Dravidian *vari*, *vari-inč* etc. This is not very attractive semantically and phonetically. Through this *ə-rig, Witzel attempts to connect the Dravidian and Indo-Aryan forms with Malay *beras* 'dehusked rice' and Writ. Tibetan *bras* 'rice'. A direct connection between these Tibetan and Malay forms and the south Asian forms, without the Munda intermediary, would be more satisfying both semantically (both sides 'rice') and phonetically (both sides /labial Consonant (+ vowel) + r-type consonant + vowel—sibilant/).

The second theory—that rice agriculture was introduced to south India by the Mundas on their westward migration from the Austroasiatic homeland in southeast Asia—is defended by archaeologists Glover & Higham (1996:419),¹⁹ Higham (2002, 2009) and Bellwood (2005): a migration in the 3rd mill. BCE would have brought the Mundas to Eastern India and, with them, rice cultivation. Blust (1998) supports that view from linguistics. He assumes a genetic relationship of proto-Austroasiatic with proto-Austronesian within "Austic", a construct which, following earlier authors (Schmidt 1906, Shorto 2006, Reid 1994) but unlike Reid (2005) he views as monophyletic. He maintains that the absence of lexical evidence for Austic is an argument in favor of its great age. He claims the Austic expansion resulted from the domestication of rice, an event he places—against archeological plausibility—in Yunnan c. 9000 BCE. A serious problem with the

¹⁹ 'In summary we can say that, towards the end of the 3rd millennium bc, rice, including domesticated varieties, appeared among the small-scale neolithic farming communities of the central and eastern parts of the Ganga valley, perhaps brought by communities of farmers speaking Proto-Munda languages expanding down the Brahmaputra valley from a homeland in the region of Yunnan and upper Burma'.

Austric-rice-expansion idea is the absence of any shared rice-specific vocabulary between proto-Austronesian and proto-Austroasiatic (Sagart 2003). A further problem specifically with the view that the PAustroasiatics introduced rice cultivation to India is the conspicuous absence of any Austroasiatic rice-specific vocabulary in south Asian languages, whether Indo-Iranian or Dravidian.

Recently Ferlus (2010) has sought to remedy the absence of Austroasiatic-Indo-Aryan matches in rice vocabulary by connecting the south Asian term *vr̥thi* etc. to a putative proto-Austroasiatic word *C.rac "rice", supposedly brought by the Mundas, with rice, to south Asia. The argument for this Austroasiatic reconstruction is very indirect and the supporting evidence very flimsy, however.

One thing is worth mentioning when one considers the Austroasiatic vocabulary of rice: the list of proto-Austroasiatic terms above —like the list earlier given by Zide et al (1976) and the more recent one by Ferlus (2010)— does not include any obvious loans from outside sources, be they east or south Asian. This situation implies a largely self-contained tradition of rice: it makes it possible that one start of rice cultivation in Asia is due to linguistics ancestors of the Austroasiatics. We will discuss this point further in our conclusion.

3.11. South Asia

McCouch and her group have documented the transfer into *indica* of domestication genes from *japonica* rice. This requires contact between peoples growing *indica* and those growing *japonica*. The Tibeto-Burmans are sino-Tibetans and members of the Sino-Tibetan-Austronesian macro-family whose earliest speakers grew *japonica* rice just north of the mid-Yangzi domestication center. After separating from the Sinitic branch of Sino-Tibetan sometime in the late 4th millennium CE in the upper Yellow River valley, the TBs expanded south and west, some of them skirting the Himalayan foothills and entering NE India, while others penetrated the Himalayan plateau in a westerly and south-westerly direction. Contact between the TBs and Indo-European peoples (Tokharians and Indo-Iranians) is probably very ancient. Today the contact front between Tibeto-Burman languages and Indo-Aryan languages in northern India is very extensive (Illustration 2):

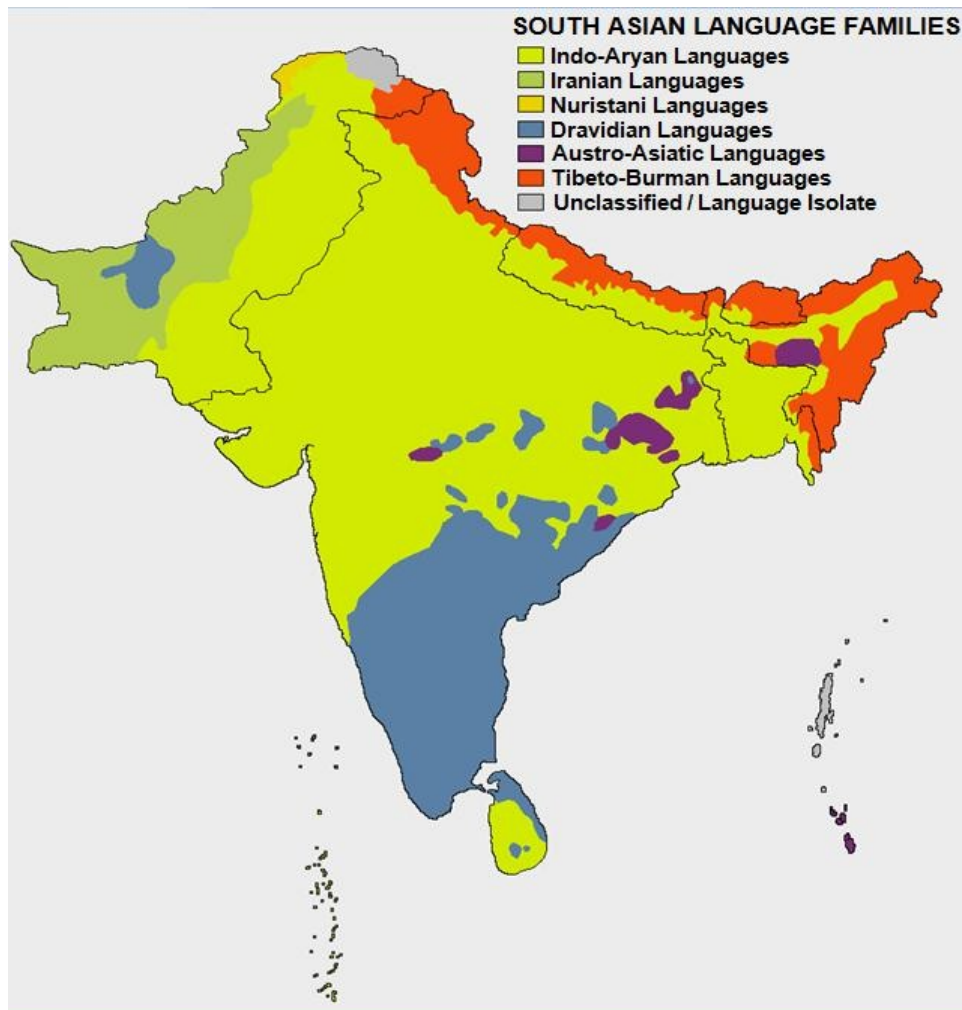


Illustration 2: The extensive contact front between Indo-Aryan and Tibeto-Burman languages

This leads one to wonder whether the domestication genes which passed into *indica* were not from *japonica* plants brought to the Himalayan region by Tibeto-Burman farmers ultimately from NW China. The phonetic resemblance between Ved. *vr̥thi*, Persian *birinj* etc. and Tibetan ^mbras 'rice', independently noted by Witzel, Ferlus and others, suggests the possibility that the south Asian forms are loans, possibly independently made, from one or several early Tibeto-Burman languages, although the phonetic path is far from clear. The possibility of an accidental resemblance cannot be excluded either.

3.12. Late transmission of indica rice varieties from India to east Asia ?

Above I have shown that Chinese textual evidence does not vindicate the idea that the *japonica/indica* distinction was known to the Chinese 2000 years ago. There is in fact evidence from language and from Chinese written sources for a late introduction of rice varieties into southeast Asia and China from India and from indianized states.

The first point is the phonetic identity between a Dravidian term for 'paddy': Tamil *vari* 'paddy', Telugu *vari* 'paddy' (Burrow 1984), on the one hand, and the Austronesian languages Malagasy *vary* 'rice' and Nadju Dayak (Borneo) *bari* 'cooked rice'. Neither Austronesian form is relatable to any of the inherited Austronesian terms for 'rice'. Adelaar (2009) shows the Malagasy migration carried Barito speakers out of south Borneo on Srivijayan Malay ships. Srivijaya was a Buddhist maritime state which existed in Sumatra from at least 671 CE to the 13th century. Its strong cultural relations with India have left a significant number of Sanskrit loans in Malay and in other Austronesian languages in the region of Sumatra-Java-Borneo. Some of these loanwords are found in Malagasy: they must have been absorbed before the migration to Madagascar. Historical sources cite trade relations between Srivijaya and the Tamil Chola dynasty in south India. It is quite likely that the Dayak-Malagasy term *bari/vary* is a Dravidian loan to certain Austronesian languages of Borneo. Borrowing of a word for 'rice' by a people already well acquainted with rice makes sense if hitherto unknown rice varieti(es) are introduced. We may have here a linguistic trace of the introduction of indica varieties from south India into insular southeast Asia.

The second point relates to China. The Chinese Song dynasty history (compiled c. 1345 CE, as cited in Zhou Shilu, 1938)²⁰ states that in 1012 or 1013 CE there was a drought in Jiangsu, Zhejiang and other regions in the Yangzi and Huai river valleys in China. The Song state sent an envoy to Fujian to arrange for large quantities of rice to be imported from the Chamic empire in central Vietnam and distributed in the drought-stricken provinces for seed. The following year, rice from the Cham empire was displayed in front of assembled provincial officials by the Song emperor. "Cham rice" was described as differing from Chinese rice in that it had long and thin grains, no awns, and was early-maturing. Zhou Shilu comments on the Song history passage, arguing that the rice obtained from the Chams must be indica rice, and that the event described in the Song history was the occasion on which indica rice was introduced to China on a large scale. He adds that modern-day peasants in Guangdong and Fujian refer to indica rice as "Cham rice".

In both the Borneo and Chinese cases, indianized states (Srivijaya, the Chamic state) played the role of intermediaries between India and East Asia.

4. Recapitulation

Let us recapitulate the main points of our discussion:

²⁰ The author thanks G. Second for bringing this text to his attention.

- there exists linguistic evidence that the Sino-Tibetans, proto-Austronesians and their linguistic ancestors the proto-Sino-Tibetan-Austronesians, cultivated rice. The proto-Sino-Tibetan-Austronesians were originally *Setaria* farmers who had acquired rice from their southern neighbours. There is however no hard evidence that specific rice-related words in proto-Sino-Tibetan-Austronesian languages were borrowed from the pre-Hmong-Miens.
- The Hmong-Mien lexicon of rice includes a majority of Chinese loanwords.
- Japanese rice agriculture may have been acquired at the contact with pre-proto-Austronesian speakers, perhaps in Shandong: the word *kome* appears to match proto-Austronesian **Semay* and related forms in the other Sino-Tibetan-Austronesian languages.
- Austronesian japonica rice was brought 'back' to the southeast Asia mainland by the Tai-Kadai; but the Tai-Kadai borrowed dry rice vocabulary and techniques from the Austroasiatics with whom they came into contact.
- The Tibeto-Burman languages have a very long front of contact with the Indo-Aryan languages, suggesting domestication genes from japonica came in rice plants grown by Tibeto-Burman speakers; there is a possibility that south Asian terms for 'rice' such as Vedic *vr̥hi* were borrowed from Tibeto-Burman.
- The Austroasiatic family has a reconstructible vocabulary indicating dry rice cultivation, without any evidence of lexical borrowing from any other source (although there are Austroasiatic loans in the Tai-Kadai rice vocabulary).
- The Indo-Aryan and Dravidian rice vocabularies include no Austroasiatic words.
- The Malagasy and Naju Dayak terms *vary* 'paddy' and *bari* 'cooked rice' are loans from Dravidian, a probable indication of the introduction of Indian rice varieties into insular southeast Asia.

Conclusion

We are now in a position to give a preliminary answer to the question in the title of this paper:

two, at least.

One is the very self-contained proto-Austroasiatic vocabulary of rice, and the other, the rest of the rice vocabularies of Asian languages, which are potentially related by vocabulary transfers.²¹ This suggests the following hypothetical model.

²¹ To the exclusion of Korean; but geography practically excludes a start of rice cultivation by Korean peoples.

The fact that proto-Austroasiatic linguistic typology is very similar to that of the East Asian groups (Sino-Tibetan-Austronesian, Hmong-Mien) implies a geographical proximity of proto-Austroasiatic with these groups (Sagart, in press). This argues against a south Asian homeland, and makes a leading role of the Austroasiatics in the domestication of *indica* unlikely. Lack of a clear Austroasiatic contribution to the rice-specific vocabulary of *indica*-cultivating peoples also argues against an association of the Austroasiatics with the start of cultivation of *indica* rice. There remains a possible association between the Austroasiatics and the domestication of *Aus* rice, the third clade in the Garris et al. dendrogram. This would be consistent with the hints of dry rice cultivation by the early Austroasiatics (above). Whether that option is open depends on plant genetics showing that *Aus* was domesticated independently from *japonica*, without introgression of *japonica* domestication genes.²² If so, *Aus* rice may have been domesticated somewhere in southeast Asia uplands by the linguistic ancestors of the Austroasiatics and brought west during the Austroasiatic migrations to India (Khasi, Munda). Any archaeological correlates are, so far, missing.

The evidence from transfers of vocabulary links the Sino-Tibetan-Austronesian macro-family (Sino-Tibetan plus Austronesian including Tai-Kadai) languages, as donors, to Japanese in the northeast and more doubtfully to Indo-Aryan in the west. Yet the presence of *Setaria italica* at the deepest level in the Sino-Tibetan-Austronesian group indicates too northerly a location for proto-Sino-Tibetan-Austronesian speakers to be behind the start(s) of *japonica* cultivation in the Yangzi valley: it is most likely that proto-Sino-Tibetan-Austronesian speakers received *japonica* rice from their southern neighbors. Judging from their geographical location and from the original rice-specific terms in their lexicon, the Hmong-Mien peoples are the most likely candidates for that role. However, hard evidence of early transfers of vocabulary from the very young Hmong-Mien language family into the earliest level of Sino-Tibetan-Austronesian has not been found. Perhaps partial relexification by Chinese in historical times is to be blamed for the fact. Alternatively, the pre-Hmong-Mien may have been onlookers at the start of *japonica* cultivation in the Yangzi valley, while the language(s) of the original *japonica* rice farmers were wiped out by the expansion of Chinese in historical times.

Geographical expansion of the Sino-Tibetan-Austronesian family is the main mechanism which explains the distribution of modern *japonica* rices: from a nuclear area located in or near Henan, the family's eastern branch brought *japonicas* (still close to the original, tropical type) to the east China coast and south to Taiwan; from there *japonicas* spread with the Austronesian peoples to insular

²² This was suggested in an oral remark by Susan McCouch, October 2010.

southeast Asia. Meanwhile, the family's western branch brought Japonicas to the upper Yellow river, then south into Sichuan, Yunnan, along the Himalayan piedmont to the edge of the Gangetic plain and ultimately to the vicinity of the Indus valley. A Tibeto-Burman arrival in northwestern India in the second millennium CE is not implausible. At the same moment, indica rices would be in the process of being domesticated; genetic exchanges between the two varieties along the long contact front shown in Illustration 2 would have led to introgression of domestication genes from the more domesticated japonica into the less domesticated indicas. Much later, in the first millennium CE, fully domesticated indica varieties would have been introduced to SE Asia through the intermediary of the indianized states (Srivijaya, Angkor, the Chamic empire etc.).

This hypothetical model of the spread of rice and languages in Asia is offered for falsification to colleagues at this symposium.

References

- Adelaar, A. (2009) Towards an integrated theory about the Indonesian migrations to Madagascar, in (P. N. Peregrine, I. Peiros and M. Feldman eds): *Ancient human migrations: a multidisciplinary approach*. 149-172. Salt Lake City: University of Utah Press.
- Anderson, G.D.S. (2004) Advances in Proto-Munda reconstruction, *Mon-Khmer Studies* vol. 34, pp. 159-184
- Bahuchet, S. (2006) Languages of African rainforest « pygmy » hunter-gatherers: language shifts without cultural admixture. Paper presented at the conference on *Historical linguistics and hunter-gatherers populations in global perspective*, Leipzig : Germany, 10-12/8/06. Online at <http://hal.archives-ouvertes.fr/hal-00548207/en/>.
- Bellwood, P. (2005) Examining the Farming/language hypothesis in the East Asian context. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds) *The Peopling of East Asia*, 17–30. London: RoutledgeCurzon.
- Benedict, P.K. (1972a) *Sino-Tibetan: a Conspectus*. Cambridge: University Printing House.
- Blench, R. (2009) If agriculture cannot be reconstructed for proto-Sinotibetan what are the consequences ? Paper presented at the 42nd Conference on Sino-Tibetan Language and Linguistics, and subsequently revised, Chiang Mai, November 2-4, 2009. Online at <http://www.rogerblench.info/Language%20data/Sino-Tibetan/Blench%20ICSTLL42%20Chiang%20Mai%20paper.pdf>

- Blench, R. and M. Post (2010) Rethinking Sino-Tibetan phylogeny from the perspective of Northeast Indian languages. Paper from the 16th Himalayan Languages Symposium, 2-5 September 2010, School of Oriental and African Studies, London.
- Blust, R. (1998) Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. In: Ward H. Goodenough (ed.) *Prehistoric settlement of the Pacific*. Philadelphia: American Philosophical Society.
- Brace, L., and T. Nagai (1982) Japanese tooth size: past and present. *American journal of physical anthropology* 59:399-411.
- Bradley, D. (1979) *Proto-Loloish*. Scandinavian Institute Monograph series N° 39. London and Malmö: Curzon.
- Bray, F. (1984) Agriculture. *Science and Civilization in China*, ed. by J. Needham, Vol. 6 part 2. Cambridge: Cambridge University Press
- Burrow, T. (1984) *A Dravidian etymological dictionary*. 2nd ed. Oxford [Oxfordshire]: Clarendon Press.
- Cauquelin, J. (1991) *Dictionnaire puyuma-français*. Paris, Jakarta: Ecole Française d'Extrême-Orient.
- Chang Kwang-chih (1980) *Shang civilization*. Newhaven and London: Yale University Press.
- Coblin, S. (1986) *A Sinologist's handlist of Sino-Tibetan Lexical Comparisons*. Monumenta Serica Monograph Series XVIII (Roman Malek, ed.). Nettetal: Steyler Verlag. 186 p.
- Diffloth, G. (2005) The contribution of linguistic palaeontology to the homeland of Austroasiatic. In L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics* 77-89. London: RoutledgeCurzon.
- Ferlus, Michel (2010) "The Austroasiatic Vocabulary for Rice: its Origin and Expansion." *Journal of the Southeast Asian Linguistics Society* 3.2 (2010), 61-76.
- Ferrell, R. (1982) *Paiwan dictionary*. Pacific Linguistics C73. Canberra, A.N.U.
- Fogg, Wayne H. (1983) Swidden cultivation of foxtail millet by Taiwan aborigines: a cultural analogue of the domestication of *Setaria italica* in China. in D. Keightley (ed.) *The origins of Chinese civilization*. Berkeley: University of California Press.
- Fu Daxiong (2001) Xizang Changguogou yizhi xin shiqi shidai nongzuowu yicun de faxian, jiangding he yanjiu. *Kaogu* 2001, 3: 66-74.
- Fuller, Dorian Q., Emma Harvey & Ling Qin (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region.

Antiquity 81 (2007): 316–331

- Garris, A.J, T. H. Tai, J. Coburn, S. Kresovich and S. McCouch (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169(3):1631-8.
- Glover I., and Higham C. (1996) New evidence for early rice cultivation in South, Southeast and East Asia. In *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, D. Harris (ed.). Smithsonian Institution Press: Washington, DC; 413–441.
- Han Kangxin and Takahiro Nakahashi (1996) A Comparative Study of Ritual Tooth Ablation in Ancient China and Japan. *Anthropol. Sci.* 104(1), 43-64.
- Haudricourt, A.-G., and D. Strecker (1991) Hmong-Mien (Miao-Yao) loans in Chinese". *T'oung Pao*, Vol. LXXVII, 4-5 (1991), pp. 335-342.
- Higham, C. (2002) Languages and farming dispersals: Austroasiatic languages and rice cultivation. In Peter Bellwood & Colin Renfrew (eds) *Examining the Farming/language dispersal hypothesis*, 223–232. Cambridge: McDonald Institute.
- (2009) East Asian Agriculture and its impact. In Chris Scarrre (ed.) *The Human past, Second edition*, 234–263. London: Thames and Hudson.
- Ino, Yoshinori (1998) Dong Ying You Ji. In: T. Moriguchi (ed.) *Ino Yoshinori Fanyu Diaocha Shouce*, pp. 205-225. Taipei: Southern Materials Center.
- Joseph, U.V. and R. Burling (2006) *The comparative phonology of the Boro Garo languages*. Mysore: Central Institute of Indian Languages.
- Kovach, Michael J., Megan T. Sweeney and Susan R. McCouch (2007) New insights into the history of rice domestication. *Trends in Genetics* 23, 11, 578-587.
- Krishnamurti, Bhadriraju (2003) *The Dravidian languages*. Cambridge: Cambridge University Press.
- Kuiper, Fransiscus B. J. (1948) *Proto-Munda words in Sanskrit*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- (1950) *An Austro-Asiatic myth in the RV*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- Li, Hui-lin (1983) The Domestication of Plants in China: Ecogeographical Considerations. In: David N. Keightley (ed.) *The origins of Chinese Civilization*, 21-64. Berkeley, Los Angeles, London: University of California Press.
- Li, Xiaoqiang, John Dodson, Zhou Xinying, Zhang Hongbin, Ryo Matsumoto (2007) Early cultivated wheat and broadening of agriculture in Neolithic China. *The Holocene* July 2007 17: 555-560.

- Londo, Jason P., Yu-Chung Chiang, Kuo-Hsiang Hung, Tzen-Yuh Chiang, and Barbara A. Schaal (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences*, June 20, 2006, vol. 103, no. 25: 9578-9583.
- Matisoff, J. (2003). *Handbook of Proto-Tibeto-Burman*. Berkeley, Los Angeles, London: University of California Press.
- Ostapirat, Weera (2000) Proto-Kra. *Linguistics of the Tibeto-Burman Area* 23.1.
- (2005) Tai-Kadai and Austronesian: notes on phonological correspondences and vocabulary distribution. In L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics* 109-133. London: RoutledgeCurzon.
- Pietrusewsky, M., A. Lauer, C.-H. Tsang (2009) Health status and lifestyle in early neolithic and iron age Taiwan. Slide show presented at the conference on *Pacific Island Archaeology in the 21st Century: Relevance and Engagement, Koror, Republic of Palau, July 1-3, 2009*.
 Online at:
http://www.anthropology.hawaii.edu/People/Faculty/Pietrusewsky/pdfs/PalauPPT_for%20website.pdf
- Pittayawat, Pittayaporn (2009) *The phonology of proto-Tai*. Unpublished Cornell U. dissertation.
- Pourrias, L. and M. Poinot (undated ms.) *Dictionnaire 'amis français*. Privately published.
- Ratliff, M. (2010) *Hmong-Mien Language History*. Canberra: Pacific Linguistics.
- Reid, Lawrence A. (1993) Another look at the language of the Tasaday. Keynote Lecture presented to the 3rd Annual Conference of the Southeast Asian Linguistic Society (SEALS III), Honolulu, Hawai'i, May 10-17, 1993.
- (1994) Morphological evidence for Austric. *Oceanic Linguistics* 33, 2: 323-344.
- (2005) The current status of Austric. A review and evaluation of the lexical and morphosyntactic evidence. In L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics* 134-160. London: RoutledgeCurzon.
- Robbeets, M. (2005). *Is Japanese Related to Korean, Tungusic, Mongolic and Turkic?* (Turcologica 64.) Wiesbaden: Harrassowitz.
- Sagart, L. (1995a) "Some remarks on the Ancestry of Chinese". In: William S.-Y. Wang (ed.) *The Ancestry of the Chinese Language*. Journal of Chinese Linguistics monograph series no. 8,

pp. 195-223.

- (1995b) "Chinese 'buy' and 'sell' and the direction of borrowings between Chinese and Miao-Yao". *T'oung Pao* LXXXI, 4-5: 328-342.
- (2003) The vocabulary of cereal cultivation and the phylogeny of East Asian languages. *Bulletin of the Indo-Pacific Prehistory Association* 23 (Taipei papers, Volume 1): 127-136.
- (2005) Sino-Tibetan-Austronesian: an updated and improved argument. In L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics* 161-176. London: RoutledgeCurzon.
- (2006) Review: James A. Matisoff (2003) Handbook of Proto-Tibeto-Burman. System and philosophy of Sino-Tibeto-Burman Reconstruction. *Diachronica* XXII, 206-223.
- (2008) the expansion of setaria farmers in east asia: a linguistic and archaeological model. In Sanchez-Mazas A, Blench R, Ross M, Peiros I, Lin M, eds. *Past human migrations in East Asia: matching archaeology, linguistics and genetics*, 133-157. Routledge Studies in the Early History of Asia, London: Routledge. http://anthro.unige.ch/~sanchez-mazas/phm/14SAGART_CHAP05-OK.pdf
- (to appear, 2011) The Austroasiatics: East to West or West to East ? in: N. J. Enfield (ed.) *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Canberra: Pacific Linguistics.
- Shorto, Harry L. (2006). *A Mon-Khmer comparative dictionary*, edited by Sidwell, Paul, Cooper, Doug and Bauer, Christian. Canberra: Australian National University. Pacific Linguistics 579.
- Sidwell, P. (2009) *Classifying the Austroasiatic Languages: history and state of the art*. Munich, Lincom Europa.
- Starostin, S.A. (2008) Altaic loans in Old Chinese. In Sanchez-Mazas A, Blench R, Ross M, Peiros I, Lin M, eds. *Past human migrations in East Asia: matching archaeology, linguistics and genetics*, 254-262. Routledge Studies in the Early History of Asia, London: Routledge.
- Schmidt, W. (1906) *Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentralasiens und Australasiens*. Braunschweig: Friedrich Vieweg und Sohn.
- Tsang, Cheng-hwa (2005) Recent discoveries at a Tapenkeng culture site in Taiwan: implications for the problem of Austronesian origins. In L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*. London: RoutledgeCurzon.

- Tsuchida, S. (1976) *Reconstruction of Proto-Tsouic phonology*. Tokyo: Study of languages and cultures of Asia and Africa monograph series N° 5.
- Unger, James M. (2008) *The Role of Contact in the Origins of the Japanese and Korean Languages*. Honolulu: University of Hawai'i Press, 2008
- van Driem, G. (1999) A new theory on the origin of Chinese. *Indo-Pacific Prehistory Association Bulletin* 18, 2:43-58.
- (2009) Rice and the Austroasiatic and Hmong-Mien homelands. Paper presented at the 4th International Conference on Austroasiatic Linguistics, Mahidol University, 29 October 2009.
- Vovin, A. (1994) Is Japanese related to Austronesian ? *Oceanic Linguistics* 33, 2: 269-390.
- Wang, T. Y. (1989): Buckwheat germplasm resources in Tibet. in: Buckwheat Research Association in China (ed.) *A collection of scientific treatises on buckwheat in China*, 49-51. Scientific Publisher, Beijing (in Chinese).
- Wang Fushi and Mao Zongwu (1995) *Miao-yao yu guyin gouni* [a reconstruction of the sound system of Proto-Miao-Yao]. Beijing: Zhongguo Shehui Kexue.
- Whitman, John Bradford (1985) *The Phonological Basis for the Comparison of Japanese and Korean*. PhD thesis, Harvard University.
- (to appear) The relationship between Japanese and Korean. in: Tranter, David N (ed.) *The Languages of Japan and Korea*. London: Routledge.
- Witzel, M. (1999) Early Sources for South Asian Substrate Languages. *Mother Tongue Special Issue*, Oct. 1999.
- (2000) *The languages of Harappa*. Online at <http://www.people.fas.harvard.edu/~Witzel/IndusLang.pdf>.
- Zhou, Shilu (1938) *Zhongguo shi dao de yuanchandi* [China is the cradle of domesticated rice]. *Zhongguo dao zuo* 7, 5: 53-54.
- Zide, A. and N. Zide (1976) Proto-Munda cultural vocabulary: evidence for early agriculture. In Philip Jenner, Laurence Thompson and Stanley Starosta (eds.) *Austroasiatic Studies part II*: 1295-1334.