

# How Many Independent Rice Vocabularies in Asia?

Laurent Sagart

Received: 30 September 2011 / Accepted: 10 December 2011 / Published online: 5 January 2012  
© Springer Science+Business Media, LLC 2011

**Abstract** The process of moving from collecting plants in the wild to cultivating and gradually domesticating them has as its linguistic corollary the formation of a specific vocabulary to designate the plants and their parts, the fields in which they are cultivated, the tools and activities required to cultivate them and the food preparations in which they enter. From this point of view, independent domestications of a plant can be expected to result in wholly independent vocabularies. Conversely, when cultivation of a plant spreads from one population to another, one expects elements of the original vocabulary to spread with cultivation practices. This paper examines the vocabularies of rice in Asian languages for evidence of linguistic transfers, concluding that there are at least two independent vocabularies of rice in Asia. This suggests at least two independent starts of cultivation and domestications of Asian rice.

**Keywords** Rice · Domestication · Asia · Linguistics

## The spread of rice as a multidisciplinary problem

How can geneticists, anthropologists, archaeologists, cultural historians, linguists, arrive together at a deeper understanding of recent human prehistory in 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 refinement, introduced in Bellwood (2005b) that while agriculture per se will normally induce an increase in population density, 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. In that sense, exclusive specialization in one kind of cereal is not conducive to wide geographical expansion.

With rice, the 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. There is no consensus on the number of starts of cultivation for rice. There are currently theories that say Asian rice, *O. sativa*, was put into cultivation only once (the “snowball model” of Vaughn et al.) and theories that say there were at least two (Kovach et al. 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. 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 three main clades: Indica, Japonica, and Aus.

To linguists, a start of cultivation is a time during which a specific vocabulary is formed in a language to attend to the

---

L. Sagart (✉)  
CRLAO,  
42bis Avenue de la Belle Gabrielle,  
75012 Paris, France  
e-mail: laurent.sagart@gmail.com

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 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.

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 and 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 linguistic ancestors 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, 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.

### 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 (a.k.a. Miao-Yao), Sino-Tibetan, Korean and Japonic. The following do not: Tungusic, Mongolic and 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.

### East Asian rice vocabularies

#### Japanese and Korean

The Japanese and Korean people cultivate temperate japonica varieties. The two languages may be genetically related (Whitman 1985, 2011), 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.

#### 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 ('Proto-Malayo-Polynesian') ancestral to all the Austronesian languages outside of Taiwan. Knowledge of rice by the proto-Austronesians is widely recognized by linguists based on three reconstructed items: proto-Austronesian \*pajay 'rice plant', \*Semay 'rice as food' and \*beRas 'husked rice'. The latter includes a monosyllabic root<sup>1</sup> \*-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.<sup>2</sup> Large quantities of carbonized rice grains were discovered in 2002–2003 at Nan Kuan Li, a lowland site on

<sup>1</sup> 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).

<sup>2</sup> In Puyuma, a language of SE Taiwan, *bəras* means 'husked grain of rice or millet' (field notes, September 12, 2002).

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 *S. italica*, \*beCeŋ, had been reconstructed to proto-Austronesian. The Nan Kuan Li site attests to co-cultivation of rice and *S. italica* by the early Austronesians on Taiwan as early as the first half of the third millennium 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, arboriculture and increased reliance on fishing presented attractive alternatives. Japonicas dominate among the traditional landraces maintained by the Austronesians in Taiwan.

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 floodlands along rivers. The root occurs as a bound morpheme in Tsou *cxana* ‘wet rice field’ (analyzable as *cx*-<proto-Austronesian \*CeNaq ‘mud’ plus \*na ‘flood-land’), 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.

#### Austronesian and Tai-Kadai

Based on shared innovations in the personal pronouns, numerals 5–10 and morphological innovations, Sagart (2004, 2005) argues that Tai-Kadai is a subgroup of Austronesian coordinate with Malayo-Polynesian, 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 2011) and Paiwan *qasal* ‘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 current knowledge (Ostapirat 2005).

The Chinese word 秈 \*sa[n]>sjen>xian1 used to mean ‘indica rice’ in modern standard Chinese matches proto-Tai \*sa:l A well and appears to be a Tai loanword into Chinese<sup>3</sup>: there is no evidence that it designated indica rice, as opposed to tropical japonica rice traded from Tai-Kadai speakers in south China. Large-scale contact between Chinese and early forms of Tai began c.2,200 years ago following the establishment in the region of present day Guangzhou of the Chinese-ruled kingdom of Zhao Tuo.

After moving ‘back’ to the mainland, the Tai-Kadais came into intimate contact with Austroasiatic-speaking populations, borrowing part of their rice vocabulary (Ferlus, p.c. to the author, 27 August 2002):

- The general word for ‘rice’, proto-Tai (Pittayaporn) \*C. qaw C is comparable with Proto-Mon-Khmer (Ferlus) \*rkoʔ/rjkoʔ ‘rice plant’.
- The word for ‘swidden, dry field’, proto-Tai \*rɔj 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-Kadais moved to the mainland carrying lowland rice agriculture with them, and acquired upland rice cultivation from their new Austroasiatic neighbours. 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

<sup>3</sup> The assertion is often made that the distinction between japonica and indica rice was known to the Chinese about 2,000 years ago, as 粳 jīng (japonica) vs. 秈 xiān (indica). While these meanings are those attached to these characters today, it is not at all certain that they designated japonica and indica varieties 2,000 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 eleventh century dictionary. It says <<江南呼粳為秈>>“Xiān is the name of Jīng 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。>>“Jīng is a kind of rice”. In a text from the Jin dynasty 265–419 CE, we learned that Jīng was dependent on irrigation: “Jīng 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 Jīng was non-sticky jīng means non-glutinous rice 稻 (玉篇: <<粳, 不黏稻>>). It would appear, therefore, that the term 秈 xiān designated nonglutinous lowland rice from south of the Yangzi.

were looking for fishing spots in river estuaries, cultivating rice on the riversides by taking advantage of seasonal flooding, as a complement to fishing. These would be coastal Formosans who did not practice cultivation of upland rice and did not have the upland landraces with them. The Austroasiatics, in contrast, were specialized in upland/swidden rice, and transmitted the technology, landraces and attendant vocabulary to the Tai-Kadais.

## 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 of these are 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 (spoken in north China in the first half of the first millennium BCE) is the earliest reconstructible form of Chinese. Introduction of non-japonica varieties in early eleventh century CE China is historically documented (below).

Not all Tibeto-Burman groups cultivate rice. Especially on the Tibetan plateau and surrounding highland regions, highland barley 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 implies knowledge of rice by the proto-Sino-Tibetans:

Old Chinese 米 \*C.m<sup>5</sup>[e]j?>mejX>mǐ ‘millet or rice grains, dehusked and polished’,<sup>4</sup> 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: rice ‘in the pot’, 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 correspondences within Sino-Tibetan are not well understood, in particular those relating to initial stop manner and to tone

(Sagart 2006); neither do we have a reliable reconstruction of proto-Sino-Tibetan which would allow us to say that Old Chinese<sup>5</sup> \*C.m<sup>5</sup>[e]j? 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 \*? match in a significant number of forms (Table 4).

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, taking into consideration of the geographical distance between these two.

The Sino-Tibetan languages share a word for *S. italica*:

稷 \*[ts]ək>tsik>ji ‘millet (*S. italica*)’  
 Lhokpu cək ‘*S. 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’.<sup>6</sup>

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 fifth 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 fifth millennium 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 fourth millennium BCE. Breakup of proto-Sino-Tibetan into a Chinese and a Tibeto-Burman branch may be dated to the late fourth millennium BCE when the western Majiayao culture c.3100–2700 BCE, in the upper Yellow river (Xishanping,

<sup>4</sup> This form is reconstructed with vowel /i/ in Baxter-Sagart v. 1.00; we now reconstruct \*e because two words with 米 as phonetic are read mjie and mjieX in MC; if it were \*-ij? we would expect MC mjijX.

<sup>5</sup> 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>.

<sup>6</sup> *S. italica* makes valuable fodder.

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

|       | Proto-Bodo-Garo | Old Chinese                         |
|-------|-----------------|-------------------------------------|
| Dream | jV 3-maŋ        | 夢 *C.məŋ-s > mjuwngH > mèng ‘dream’ |
| Lose  | Gw 1-ma         | 無 *ma > mju > wú ‘not have’         |
| Name  | muŋ 1           | 名 *C.meŋ > mjieng > míng ‘name’     |

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<sup>h</sup>iŋ>den>tián ‘field; to hunt’. Outside of written Tibetan the term also occurs in Lepcha *lyǎŋ*, Cuona *lej*<sup>13</sup>, Hayu *jing* ‘dry field’, although some of these forms could be borrowed from Tibetan.

Objections to a North China Homeland for Sino-Tibetan

Identification of proto-Sino-Tibetan with a stage of Yangshao culture 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 specialized 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.

Starostin (2008) envisioned a Sino-Tibetan homeland in the Himalayan region; he assumed a Chinese migration east out of the Sino-Tibetan homeland after 3000 BCE. He equated the archaeological Yangshao culture with proto-Altaiic (a 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

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

|             | Proto-Bodo-Garo      | Old Chinese                              |
|-------------|----------------------|--|
| Change      | *lai (no tone given) | 易 *lek > yek > yi ‘change; exchange’     |
| Spirit, god | *mw-Dai 4            | 帝 *t <sup>h</sup> ek-s > tejH > di ‘God’ |

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

|       | Proto-Bodo-Garo | Old Chinese                                     |
|-------|-----------------|---|
| Die   | thui 1          | 死 *sijʔ > sijX > sǐ ‘die (v.)’                  |
| Buy   | prai 1          | 買 *m <sup>h</sup> rajʔ > meax > mǎi ‘buy’       |
| Break | phai 2          | 破 *p <sup>h</sup> aj-s > phaH > pò ‘break (v.)’ |

reconciled with the absence of corresponding terms in proto-Altaiic: a more attractive match for proto-Altaiic 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’ views seem sound.

Starostin held that the pre-Chinese acquired agriculture at the end of their eastward migration: millet from the Altaic peoples and 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 *S. italica*<sup>7</sup>: 稷 \*[ts]ək. The source of the borrowing according to him is the Proto-Altaiic reconstruction \*žǐúgi ‘*Panicum miliaceum*’ in Starostin et al. (2003:1547-8). This reconstruction is based on proto-Turkic \*yüür ‘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).<sup>8</sup> 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).

Other opponents of a north China homeland argue that Sino-Tibetan phylogeny is poorly understood and that a primary split between Chinese and Tibeto-Burman 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 proto-language. It

<sup>7</sup> Glossed by him as ‘*Panicum*’ apparently based on Li (1983:29) against the judgment of Bray (1984), Chang (1980:147) and Fogg (1983).

<sup>8</sup> As Robbeets explains, she expects proto-Altaiic \*ž to be reflected as t- in Turkic, not y-; by her correspondences, no proto-Altaiic 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.

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

|           | Proto-Bodo-Garo | Old Chinese                               |
|-----------|-----------------|---|
| Die       | thui 1          | 死 *sijʔ > sijX > si 'die (v.)'            |
| Buy       | prai 1          | 買 *m <sup>1</sup> 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'  |

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 w<sup>1</sup>e[n]ʔ 'dog') by the earlier word for 'puppy', something like #kur/kuy (Proto-Tibeto-Burman \*kwi, 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': Proto-Tibeto-Burman (Benedict 1972) \*niy, Old Chinese \*nit; 'varnish': PTB (Benedict 1972) \*tsiy, Old Chinese \*ts<sup>h</sup>it; 'arrow': Proto-Tibeto-Burman (Benedict 1972) \*b-la, Old Chinese 弋 \*lək; 'change': PTB (Benedict 1972) \*lay, Old Chinese 易 \*lek, etc.<sup>9</sup> The low level of diversity in present day north China is sufficiently explained by the levelling role of Chinese<sup>10</sup>; 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).

### 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.

<sup>9</sup> 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.

<sup>10</sup> Levelling of linguistic diversity by northern Chinese even applies to Chinese dialects, more diverse in southeast China, despite the fact that the historical cradle of Chinese is in north China.

This claim is predicated on the observation of sound correspondences on basic vocabulary and of morphological parallels with cognate markers. The sound correspondences obtain primarily 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):

The comparisons in Table 5 obey the sound correspondences in Sagart (2005).<sup>11</sup>

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 primarily 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 who had already developed religious practices centred 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.

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 requirements from *Setaria* was instrumental in accelerating demographic growth and geographical expansion; domesticated pigs (possibly 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-ɹ<sup>1</sup>at>lat>li 'dehusked but not polished grain' and Tibeto-Burman words like Written

<sup>11</sup> Even though Sagart (2005) gives the Old Chinese vowel as \*i.

**Table 5** Two agricultural terms shared by Old Chinese and proto-Austronesian

|                            | Old Chinese             | Proto-Austronesian |
|----------------------------|-------------------------|--------------------|
| Rice, ready to cook/cooked | *C.m <sup>5</sup> [e]j? | Semay              |
| <i>S. italica</i>          | *tsək                   | beCen              |

Tibetan <sup>m</sup>bras ‘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.

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 and Qin (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 fourth 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.

#### 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 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. My own placement of the pre-Austronesians in south Shandong’s Dawenkou culture (Sagart 1995a) makes them close neighbours 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).<sup>12</sup> 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 \*kome<sub>2</sub>, proto-Japonic \*kəmai or \*kəmāi. Neither of these proto-Japonic forms looks like native Japanese words, the first because it would violate Arisaka’s laws on co-occurrence of vowels, and the second because one would expect to find komo- alternating with *kome* in modern compound words. 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.<sup>13</sup>

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-Mieng 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

<sup>12</sup> “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)

<sup>13</sup> 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*.

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.

Unlike Sino-Tibetan, Austronesian or Austroasiatic, the Hmong-Mien family is very young: any sister languages it once had are extinct. 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 therefore part of the Hmong-Mien proto-language, show phonological features which are characteristic of late Old Chinese, roughly 500–220 BCE,<sup>14</sup> a period of strong cultural dominance of Chinese in the region. Some of these words relate to technical innovations of the same period.<sup>15</sup> 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.

#### Hmong-Mien rice vocabulary borrowed into Chinese?

Haudricourt and Strecker (1991), Ratliff (2010) and 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’; 粿 ‘bread, pastry’; and 買 ‘buy’ and 賣 ‘sell’. Sagart (1995a, b) showed the 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 [yang (=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’.<sup>16</sup> In Hmong-Mien, in contrast, Ratliff (2010) reconstructs two phonetically similar but historically non-cognate forms in proto-Hmongic and proto-Mienic: \*ʔjɛŋ A and \*ʔjaŋ

A, respectively: this makes it likely that Hmongic and Mienic borrowed the Chinese term separately, in the first millennium CE. With 稻 \*[]<sup>h</sup>uʔ, whose modern form *dào* designates the rice plant in modern Chinese, the oldest tokens of the corresponding character in the Zhou bronze inscriptions (first millennium BCE) are written with the signific 米 ‘dehusked grains’, instead of 禾 ‘cereal plant’. 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 昏 \*[]u, \*[]uʔ, []<sup>h</sup>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 ‘dehusked rice grains out of the mortar’. 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 \*l̥. One of these is Old Chinese \*m.l̥, which means that one possible Old Chinese reconstruction of 稻 is \*m.l̥uʔ. This may support the idea of a connexion with the proto-Hmong-Mien word \*mbləu A ‘rice plant’.<sup>17</sup> 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 媼 \*ʔ<sup>h</sup>uʔ ‘old woman’ corresponding to Proto-Hmong-Mien \*ʔ uX ‘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 meaning as ‘bag/pocket’. She compares this with 囊 \*n<sup>h</sup>aŋ>nang>náng ‘sack, bag’, proposing that 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 lookalikes, 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 \*l̥iŋ ‘field’, 鎌 \*[]em ‘sickle’,

<sup>14</sup> 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>.

<sup>15</sup> 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).

<sup>16</sup> 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’.

<sup>17</sup> The velar cluster in Tokharian klu ‘rice’, which is usually regarded as a loan from Chinese, may be an attempt to reproduce the initial cluster in late Old Chinese form of this word: [l̥<sup>h</sup>uʔ]: at least this is the strategy used by Hmong-Mien, compare Old Chinese 桃 \*C.l̥aw>-daw>táo ‘peach’, Proto-Hmong-Mien \*gləw A ‘peach’ (Ratliff 2010). Note that Tokharian does not possess the cluster \*gl-.



糗 \*[g](r)a(k)-s ‘cakes’, 擣 \*t<sup>h</sup>u? ‘pound, thresh’ and 甑 \*s-təŋ-s ‘steamer’ are loans from Hmong-Mien. This cannot be true of ‘sickle’: if proto-Hmong-Mien \*ljim ‘sickle’ had been borrowed by Chinese before the fourth 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 first century CE and would appear in Middle Chinese with initial y-, not l-. It is also demonstrably false for ‘steamer’, 甑 \*s-təŋ-s > tsiŋH > zèng, 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 糝 \*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: 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. The source of the borrowing is 食 \*mə-lək > zyik > shí ‘eat’, the regular Chinese word in this meaning, a word moreover with a good Sino-Tibetan etymology: proto-Loloish \*m-lyak L ‘to lick’ (Bradley 1979, item #630). 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 they are not rice-specific terms, which weakens the argument that they are loans from Hmong-Mien. ‘Field’ moreover has Tibeto-Burman cognates (above).

#### 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 Indic, 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.

Donegan and Stampe (2004) 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, and the perceived diversity of Munda, leads them to doubt that the Austroasiatic homeland was in Southeast Asia. Even then, language contact would be useful in providing a motivation for the initial shift in speech rhythm. Sagart (2011) 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 bran’ #phe:ʔ has a certain resemblance with proto-Hmong-Mien \*mphjɛk ‘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.

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 \*cɛh, \*ha:l and \*sa:, all ‘raw rice, paddy’. None of these appear to have been borrowed from an outside source, or to have been loaned to an outside group.

Knowledge of rice by the Proto-Austroasiatics naturally raises the question whether rice cultivation in south Asia

**Table 6** Proto-Austroasiatic rice-specific vocabulary (Diffloth 2005)

|                 |           |
|-----------------|-----------|
| Rice plant      | #(kə)ʃa:ʔ |
| Rice grain      | #rəŋko:ʔ  |
| Rice outer husk | #cəŋka:m  |
| Rice inner husk | #kəndək   |
| Rice bran       | #phe:ʔ    |
| Dibbling stick  | #jərmuəl  |

ultimately goes back to a start of rice cultivation by Austroasiatic-speaking peoples. This could have happened in two ways: under the hypothesis of a south Asian homeland, the Proto-Austroasiatics 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 apparent 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 or an extinct language. A Tibeto-Burman presence in northwest India at the time of the Rgveda should not at all be deemed impossible. 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 Indic forms with Malay *beras* 'dehusked rice' and Writ. Tibetan <sup>m</sup>bras 'rice'. A direct connection between these Tibetan and Austronesian forms and the south Asian forms, without the Munda intermediary, would be more satisfying both semantically (both sides 'rice') and phonetically (both sides b/v—r—s/z).

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 and Higham (1996:419),<sup>18</sup> Higham (2002, 2009) and Bellwood (2005a): a migration in the third millennium 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 Austric is an argument in favor of its great age. He claims the Austric expansion resulted from the domestication of rice, an event he places—against the silence of archeology—in Yunnan c.9000 BCE. A serious problem with the Austric rice expansion idea is the absence of any shared rice-specific vocabulary between proto-Austronesian and proto-Austroasiatic (Sagart 2003). Pace (Sagart 2008), a further problem specifically with the view that the Proto-Austroasiatics introduced rice cultivation to India, is the conspicuous absence of any Austroasiatic rice-specific vocabulary in south Asian languages, be they Indo-Iranian or Dravidian.

Recently Ferlus (2010) has sought to remedy the absence of Austroasiatic-Indic matches in rice vocabulary by connecting the south Asian term *vr̥hi* 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 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 and Zide (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 a start of rice cultivation in Asia is due to linguistics ancestors of the Austroasiatics. We will discuss this point further in our conclusion.

## South Asia

McCouch and her group have documented the transfer into indica of domestication genes from japonica rice (Kovach et al. 2007). This requires contact between peoples growing indica and those growing japonica. The Tibeto-Burmans are 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 fourth millennium CE in the upper Yellow River valley, the Tibeto-Burmans expanded south and west, some of them skirting the Himalayan foothills and entering northeastern India, while others penetrated the Himalayan plateau in a westerly and southwesterly direction. Contact between the Tibeto-Burman and Indo-European peoples (Tokharians and Indo-Iranians) is probably very ancient. Today the contact front between the Tibeto-Burman and Indic languages is very extensive: it extends along an arc of more than 1,500 km, from northeastern Pakistan in the west to Arunachal Pradesh in the east and from Arunachal Pradesh south to Bangladesh. This leads one to wonder whether the domestication genes which passed into indica were not from japonica plants

<sup>18</sup> "In summary we can say that, towards the end of the third 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".

brought to the Himalayan region by Tibeto-Burman farmers ultimately from northwest China. The phonetic resemblance between Ved. *vr̥thi*, Persian *birinj* etc. and Tibetan <sup>m</sup>bras ‘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.

#### Late transmission of Indica Rice varieties from India to East Asia

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’ is noteworthy. 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 thirteenth 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 varietal(es) are introduced. We may have here one linguistic trace of the introduction of indica varieties from India 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 essentially 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.<sup>19</sup> This suggests the following hypothetical model.

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 with these groups (Sagart 2011). 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. This would be consistent with the hints of dry rice cultivation by the early Austroasiatics. Whether that option is open depends on plant genetics showing that Aus was domesticated without introgression of japonica domestication genes.<sup>20</sup> If so, Aus rice may have been domesticated somewhere in Southeast Asian uplands by the linguistic ancestors of the Austroasiatics and brought west during the Austroasiatic migrations to India (Khasi, Munda). 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 less certainly to Indic in the west. Yet the presence of *S. 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 neighbours. Judging from their geographical location and from the original rice-specific terms in their lexicon, the Hmong-Mien people 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 explains much of the distribution of modern japonica rice: 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 from there south to Taiwan and on to insular 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 rice would be in

<sup>19</sup> Geography practically excludes a start of rice cultivation by Korean peoples.

<sup>20</sup> This was suggested in an oral remark by Susan McCouch, October 2010

the process of being domesticated; genetic exchanges between the two varieties along the long contact front between the Tibeto-Burman and Indic languages would have led to introgression of domestication genes from the more domesticated japonicas 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 etc.).

## References

- Adelaar A. Towards an integrated theory about the Indonesian migrations to Madagascar. In: Peregrine PN, Peiros I, Feldman M, editors. Ancient human migrations: a multidisciplinary approach. Salt Lake City: University of Utah Press; 2009. p. 149–72.
- Anderson GDS. Advances in Proto-Munda reconstruction. *Mon-Khmer Studies*. 2004;34:159–84.
- Bahuchet S. 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 August 2006. 2006. <http://hal.archives-ouvertes.fr/hal-00548207/en/>.
- Bellwood P. Examining the farming/language hypothesis in the East Asian context. In: Sagart L, Blench R, Sanchez-Mazas A, editors. The peopling of East Asia. London: RoutledgeCurzon; 2005a. p. 17–30.
- Bellwood P. First farmers: the origin of agricultural societies. London: Blackwell; 2005b.
- Benedict PK. Sino-Tibetan: a Conspectus. Cambridge: University Printing House; 1972.
- Blench R. 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. 2009 <http://www.rogerblench.info/Language%20data/Sino-Tibetan/Blench%20CSTLL42%20Chiang%20Mai%20paper.pdf>.
- Blench R and Post M. 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; 2010.
- Blust R. Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology. In: Goodenough WH, editor. Prehistoric settlement of the Pacific. Philadelphia: American Philosophical Society; 1998.
- Brace L, Nagai T. Japanese tooth size: past and present. *Am J Phys Anthropol*. 1982;59:399–411.
- Bradley D. Proto-Loloish. Scandinavian Institute Monograph series N° 39. London: Curzon; 1979.
- Bray F. Agriculture. In: Needham J, Needham J, editors. Science and civilization in China, vol. 6 part 2, vol. 6. Cambridge: Cambridge University Press; 1984.
- Burrow T. A Dravidian etymological dictionary. 2nd ed. Oxford: Clarendon; 1984.
- Cauquelin J. Dictionnaire puyuma-français. Paris: Ecole Française d'Extrême-Orient; 1991.
- Chang KC. Shang civilization. Newhaven: Yale University Press; 1980.
- Coblin S. A sinologist's handlist of Sino-Tibetan lexical comparisons. In: Malek R, editor. Monumenta Serica Monograph Series XVIII. Nettetal: Steyler Verlag; 1986. 186 p.
- Diffloth G. The contribution of linguistic palaeontology to the homeland of Austroasiatic. In: Sagart L, Blench R, Sanchez-Mazas A, editors. The peopling of East Asia: putting together archaeology, linguistics and genetics. London: RoutledgeCurzon; 2005. p. 77–89.
- Donegan P, Stampe D. Rhythm and the synthetic drift of Munda. The yearbook of South Asian languages and linguistics. Berlin: De Gruyter; 2004. p. 3–36.
- Ferlus M. The Austroasiatic vocabulary for rice: its origin and expansion. *Journal of the Southeast Asian Linguistics Society*. 2010;3(2):61–76.
- Ferrell R. Paiwan dictionary. Pacific Linguistics C73. Canberra: A.N. U.; 1982.
- Fogg WH. Swidden cultivation of foxtail millet by Taiwan aborigines: a cultural analogue of the domestication of *Setaria italica* in China. In: Keightley D, editor. The origins of Chinese civilization. Berkeley: University of California Press; 1983.
- Fuller DQ, Harvey E, Qin L. Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity*. 2007;81(2007):316–31.
- Fuller DQ, Qin L. Evidence for a late onset of agriculture in the Lower Yangtze region and challenges for an archaeobotany of rice. In: Sanchez-Mazas A, Blench R, Ross M, Peyros I, Lin M, editors. *Past human migrations in East Asia*. Routledge studies in the Early History of East Asia. London: Routledge; 2008. p 40–83
- Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S. Genetic structure and diversity in *Oryza sativa* L. *Genetics*. 2005;169(3):1631–8.
- Glover I, Higham C. New evidence for early rice cultivation in South, Southeast and East Asia. In: Harris D, editor. The origins and spread of agriculture and pastoralism in Eurasia. Washington: Smithsonian Institution Press; 1996. p. 413–41.
- Han KX, Nakahashi T. A comparative study of ritual tooth ablation in Ancient China and Japan. *Anthropol Sci*. 1996;104(1):43–64.
- Haudricourt AG, Strecker D. Hmong-Mien (Miao-Yao) loans in Chinese. *T'oung Pao*. 1991; 77, 4–5, pp. 335–342.
- Higham C. Languages and farming dispersals: Austroasiatic languages and rice cultivation. In: Bellwood P, Renfrew C, editors. Examining the farming/language dispersal hypothesis. Cambridge: McDonald Institute; 2002. p. 223–32.
- Higham C. East Asian agriculture and its impact. In: Scarre C, editor. The human past. 2nd ed. London: Thames and Hudson; 2009. p. 234–63.
- Ino Y, Dong Ying You Ji. In: Moriguchi T, editor. Ino Yoshinori Fanyu Diaocha Shouce. Taipei: Southern Materials Center; 1998. p. 205–25.
- Joseph UV, Burling R. The comparative phonology of the Boro Garo languages. Mysore: Central Institute of Indian Languages; 2006.
- Kovach MJ, Sweeney MT, McCouch SR. New insights into the history of rice domestication. *Trends Genet*. 2007;23(11):578–87.
- Krishnamurti B. The Dravidian languages. Cambridge: Cambridge University Press; 2003.
- Kuiper FBJ. Proto-Munda words in Sanskrit. Amsterdam: Noord-Hollandische Uitgevers Maatschappij; 1948.
- Kuiper FBJ. An Austro-Asiatic myth in the RV. Amsterdam: Noord-Hollandische Uitgevers Maatschappij; 1950.
- Li H-L. The domestication of plants in China: ecogeographical considerations. In: Keightley DN, editor. The origins of Chinese civilization. Berkeley: University of California Press; 1983. p. 21–64.
- Li X, Dodson J, Zhou XY, Zhang HB, Matsumoto R. Early cultivated wheat and broadening of agriculture in Neolithic China. *Holocene*. 2007;2007(17):555–60.
- Londo JP, Chiang Y-C, Hung K-H, Chiang T-Y, Schaal BA. 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. 2006;103(25): 9578–9583.
- Ostapirat W. Proto-Kra. *Linguistics of the Tibeto-Burman area*. 2000;23(1):1–251.
- Ostapirat W. Tai-Kadai and Austronesian: notes on phonological correspondences and vocabulary distribution. In: Sagart L, Blench R, Sanchez-Mazas A, editors. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: Routledge Curzon; 2005. p. 109–33.
- Pietrusewsky M, Lauer A, Tsang C-H. 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. 2009. [http://www.anthropology.hawaii.edu/People/Faculty/Pietrusewsky/pdfs/PalauPPT\\_for%20website.pdf](http://www.anthropology.hawaii.edu/People/Faculty/Pietrusewsky/pdfs/PalauPPT_for%20website.pdf).
- Pittayaporn P. The phonology of proto-Tai. Unpublished Cornell University dissertation; 2009
- Pourrias L, Poinsot M. *Dictionnaire 'amis français*. Privately published; 2011
- Ratliff M. Hmong-Mien language history. Canberra: Pacific Linguistics; 2010.
- Reid LA. Morphological evidence for Austric. *Oceanic Linguistics*. 1994;33(2):323–44.
- Reid LA. The current status of Austric. A review and evaluation of the lexical and morphosyntactic evidence. In: Sagart L, Blench R, Sanchez-Mazas A, editors. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: Routledge Curzon; 2005. p. 134–60.
- Reid LA. 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.
- Robbeets M. Is Japanese related to Korean, Tungusic, Mongolic and Turkic? (*Turcologica* 64). Wiesbaden: Harrassowitz; 2005.
- Sagart L. Some remarks on the Ancestry of Chinese. In: Wang WS-Y, editor. *The ancestry of the Chinese language*. J Chin Ling. 1995a; monograph series no. 8, pp. 195–223.
- Sagart L. Chinese 'buy' and 'sell' and the direction of borrowings between Chinese and Miao-Yao. *T'oung Pao* LXXXI. 1995;4–5:328–42.
- Sagart L. The vocabulary of cereal cultivation and the phylogeny of East Asian languages. *Bulletin of the Indo-Pacific Prehistory Association* 23. 2003;1:127–36. Taipei papers.
- Sagart L. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics*. 2004;43(2):411–444.
- Sagart L. Sino-Tibetan-Austronesian: an updated and improved argument. In: Sagart L, Blench R, Sanchez-Mazas A, editors. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: RoutledgeCurzon; 2005. p. 161–76.
- Sagart L. 2006. Review: James A. Matisoff (2003) *Handbook of Proto-Tibeto-Burman. System and philosophy of Sino-Tibeto-Burman Reconstruction*. *Diachronica* XXII, 206–223.
- Sagart L. 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, editors. *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](http://anthro.unige.ch/~sanchez-mazas/phm/14SAGART_CHAP05-OK.pdf).
- Sagart L. The Austroasiatics: east to west or west to east? In: Enfield NJ, editor. *Dynamics of human diversity: the case of mainland Southeast Asia*. Canberra: Pacific Linguistics; 2011. p. 345–59.
- Schmidt W. *Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens*. Braunschweig: Friedrich Vieweg und Sohn; 1906.
- Shorto H, Sidwell P, Cooper D, Bauer C, editors. *A Mon-Khmer comparative dictionary*. Sidwell Canberra: Australian National University. *Pacific Linguistics* 579; 2006.
- Sidwell P. *Classifying the Austroasiatic languages: history and state of the art*. *Lincom Europa*: Munich; 2009.
- Starostin SA. Altaic loans in Old Chinese. In: Sanchez-Mazas A, Blench R, Ross M, Peiros I, Lin M, editors. *Past human migrations in East Asia: matching archaeology, linguistics and genetics*. London: Routledge Studies in the Early History of Asia; 2008. p. 254–62.
- Starostin SA, Dybo A, Mudrak O. *Etymological dictionary of the Altaic languages*. Leiden: Brill; 2003.
- Tsang C-H. Recent discoveries at a Tapenkeng culture site in Taiwan: implications for the problem of Austronesian origins. In: Sagart L, Blench R, Sanchez-Mazas A, editors. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: RoutledgeCurzon; 2005.
- Tsushima, S. 1976. *Reconstruction of Proto-Tsouic phonology*. Tokyo: Study of languages and cultures of Asia and Africa monograph series N° 5.
- Unger JM. *The role of contact in the origins of the Japanese and Korean Languages*. Honolulu: University of Hawai'i Press; 2008. p. 2008.
- van Driem G. A new theory on the origin of Chinese. *Indo-Pacific Prehistory Association Bulletin*. 1999;18(2):43–58.
- van Driem G. 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. Is Japanese related to Austronesian? *Oceanic Linguistics*. 1994;33(2):269–390.
- Wang TY (1989). Buckwheat germplasm resources in Tibet. In: Buck wheat Research Association in China, editor. *A collection of scientific treatises on buckwheat in China*, 49–51. Scientific Publisher, Beijing (in Chinese).
- Wang F, Mao Z. *Miao-yao yu guyin gouni [a reconstruction of the sound system of Proto-Miao-Yao]*. Beijing: Zhongguo Shehui Kexue; 1995.
- Whitman JB. The relationship between Japanese and Korean. In: Tranter, David N, editors. *The Languages of Japan and Korea*. London: Routledge; 2011
- Whitman JB. *The Phonological Basis for the Comparison of Japanese and Korean*. PhD thesis, Harvard University; 1985.
- Witzel M. Early sources for South Asian substrate languages. *Mother Tongue special issue*, Oct 1999; 1999.
- Witzel M. *The languages of Harappa*. 2000. <http://www.people.fas.harvard.edu/~Witzel/IndusLang.pdf>. Accessed September 1, 2011.
- Wolff JU. *Proto-Austronesian phonology with glossary*, vol. 2. Ithaca: Cornell Southeast Asia Program Publications; 2010.
- Zide A and Zide N. Proto-Munda cultural vocabulary: evidence for early agriculture. In: Philip Jenner, Laurence Thompson, Stanley Starosta, editors. *Austroasiatic studies part II*. 1976; 1295–334.