Historical linguistics and archaeology have a proven record of useful collaborative work. However, the distinctively different nature of the raw material of each discipline can easily lead researchers in the two disciplines toward different conclusions. It is important to recognise that this is perhaps inevitable, but that divergence between the conclusions of the two fields is sometimes the result of over interpretation of negative evidence, a methodological flaw that lies behind some archaeological arguments relating to the prehistory of Island Southeast Asia. Both linguistic and archaeological records will always be fragmentary, and where inferences drawn from them differ dramatically from one another, the most rational evaluation metric is the strength of each argument in terms of how successfully it competes against alternative explanations within its own discipline, not a prioritisation of the evidence of one academic discipline as against another.

Introduction

The prospects for an alliance between archaeology and historical linguistics that might strengthen the inferential basis of both disciplines were first explored at least 40 years ago in a special issue of *World Archaeology* dedicated to the topic ‘Archaeology and Linguistics’ (WA 8.1, June 1976). Various of the contributions to that volume showed that each discipline has resources that can illuminate aspects of the human cultural past that the other discipline lacks. In particular, archaeology can provide greater detail about material culture and settlement pattern, and has reliable methods that are not generally available to linguists for absolute dating and determining the location of prehistoric language communities. However, although it offers the possibility of inferences about non-material culture, these are normally indirect and less secure than those based on the analysis of physical artefacts. The inferences of historical linguistics, on the other hand, are not limited to those aspects of material culture that survive into the present. By identifying cognate terms with unambiguous referents, historical linguists can reconstruct a record not only of the kinds of artefacts described by archaeologists, but also of items of material culture that fail to survive in typical archaeological contexts, such as basketry, matting, traps made of wood or bamboo, and the like. In addition, the entire realm of non-material culture, including systems of kinship, marriage rules, social organisation, and ideas about the spirit world are accessible to inferences based on cognate vocabulary, but lie beyond the pale of confident archaeological errors.
investigation. For a brilliant exemplification of what this approach can contribute to prehistory, in many cases filling lacunae in the archaeological record, see the landmark multivolume study *The Lexicon of Proto Oceanic*, of which five volumes have now appeared (Ross et al. 1998, 2003, 2008, 2011, 2016).

The purpose of this paper is to remind scholars in both disciplines of what can be gained by a deepened appreciation of what each has to offer the other, but at the same time to stress that the need to remain faithful to the material of one’s own discipline can make this alliance an uneasy one, as will be seen in the following pages.

**Converging on Taiwan**

One of the great successes of convergent lines of inference originating in historical linguistics and archaeology has been the broad acceptance of the view that the movement of Austronesian (AN) languages and the related Neolithic cultures of insular Southeast Asia and the Pacific began in Taiwan, a position known for a time as the ‘Bellwood-Blust hypothesis’, but now more appropriately called the ‘Out-of-Taiwan’ (OOT) hypothesis. It is important to stress that the OOT hypothesis developed independently in each discipline. Following an earlier study (Blust 1977) in which the OOT hypothesis was implied but not stated explicitly, Blust (1984/85) argued that the AN homeland was on or near Taiwan (but not in the Philippines). At about the same time, Bellwood concluded on the basis of the earliest appearance of Neolithic cultures in insular Southeast Asia that ‘the ultimate region of Austronesian origin lay in the Neolithic landscape of southern China’ (1984/85: 109) with a subsequent settlement in Taiwan. In a letter written to me early in 1986, he expressed his interest that I had arrived at the same conclusion regarding the AN homeland toward which the archaeology was then driving him. Both linguistic and archaeological evidence for this claim has continued to accumulate over the 30 years since it was first put forth, and the OOT hypothesis is now all but universally accepted by linguists and archaeologists familiar with the relevant data. It is noteworthy that over the same time span no such unanimity of inference has appeared in the work of population geneticists, who have been split if not evenly then at least significantly, in either supporting or rejecting this view of AN origins.

**The early invisibility of rice**

Once the OOT was widely adopted as a working hypothesis, the first substantive issue that divided linguists from archaeologists was the antiquity of cereal cultivation in the AN world. As noted by Bellwood:

> remains of rice and millet were universally absent from sites of the Dabenkeng phase in Taiwan (3,500–2,500 BC) until both were found in unprecedented carbonised quantities dating to c. 2,800 BC, in hitherto unique waterlogged conditions, in the Nanguanli sites in the Tainan Science-Based Industrial Park (Tsang 2005; Tsang et al. 2004) (2011: 98).

What is striking, and one might say revealing, about this reversal relating to a major component of the archaeological record is that linguists had reported *overwhelming* evidence for rice and millet in Taiwan for more than three decades before archaeological confirmation was unearthed during Tsang’s excavations from September 2002 to March 2003. Blust, for example, who located Proto-Austronesian on Taiwan, observed that ‘[t]he linguistic evidence for cultivation of grain crops by PAN times is abundant’ (1995: 460) and supported this statement with the
suite of reconstructed terms shown in (1), where the geographical distribution of reflexes in the modern languages over major regions is signalled by T = Taiwan, P = Philippines, WIN = western Indonesia, and EIN = eastern Indonesia:

(1)

PAN term
1. *pajay 'riceplant, rice in the field’ (T, P, WIN, EIN)
2. *beRas 'harvested, husked rice’ (T, P, WIN, EIN)
3. *Semay 'cooked rice’ (T, P, WIN)
4. *baCaR 'millet sp.’ (T, P, WIN, EIN)
5. *beCeŋ 'millet, probably foxtail millet’ (T, P, WIN, EIN)
6. *zawa 'millet sp.’ (T, P, WIN)
7. *bineSiq 'seed rice’ (T, P, WIN, EIN)
8. *buRaw 'chase animals from the fields’ (T, P, WIN, EIN)
9. *qani 'to harvest, usually rice’ (T, P, WIN)
10. *lepaw 'granary’ (T, P, WIN, EIN)
11. *eRik/iRik 'thresh by trampling’ (T, WIN)
12. *paspas 'thresh by beating’ (T, P)
13. *lesuŋ 'mortar’ (T, P, WIN, EIN)
14. *qaSelu 'pestle’ (T, P, WIN, EIN)
15. *bayu 'to pound rice’ (T, P, WIN, EIN)
16. *tapeS 'to winnow grains’ (T, P, WIN)
17. *qeCah 'rice husk’ (T, P, WIN)
18. *zaRami 'rice straw, stalks left in the field’ (T, P, WIN)

It would be difficult to overemphasise the conclusiveness of this body of evidence. Whether the archaeological record showed it or not, it was obvious to well-informed AN comparativists from at least the late 1960s that speakers of PAN cultivated both rice and millet. To deny this conclusion because of negative evidence from a sister discipline would have been tantamount to rejecting the entire theoretical and methodological foundations of historical linguistics as it had developed since the first two decades of the nineteenth century, in effect denying the validity of an entire academic discipline. This is not an exaggeration: straightforward application of the comparative method that had been developed in the discovery of the Indo-European language family and tested successfully for nearly two centuries demanded the reconstruction of multiple terms for rice and millet in order to account for the wide geographical distribution of cognate forms. Then, with the excavation of Nanguanli the previously 'invisible' rice that clearly was

---

2 A decade earlier (Blust 1984/85: 61–62) nine of these 18 terms were cited as evidence for the same conclusion, namely that the AN homeland almost certainly included Taiwan, and that PAN speakers were sedentary cultivators of grain, root and tree crops, while continuing at the same time to hunt, fish, and gather wild foods. Nearly a decade before that (Blust 1976: 31–33), four of these terms, including the key three terms for rice (*pajay, *beRas, *Semay), were cited as evidence for PAN rice and millet agriculture, and although a Formosan homeland was not explicitly claimed it was implied in much of the discussion, where all reconstructions assigned to PAN required reflexes in at least one language from two or more of the following groups: Northern (Formosan), Western (insular Southeast Asian), and Eastern (Oceanic). Even earlier, Ferrell had noted that '[t]erms for rice in various stages (plant, unhusked grain, etc.) in Formosan languages are clearly AN, which should not have been the case if they had learned rice cultivation from Chinese’ (1969: 10). In the vocabularies that accompanied his text, items relating to grain agriculture that were cognate with terms already reconstructed for 'Uraustronesisch' by Dempwolff (1934–1938) included: 1. rice (plant), 2. rice (grain), 3. rice (cooked), 4. millet, 5. mortar, and 6. pestle. Full supporting evidence for the reconstructions under (1) is given in the open-access online Austronesian Comparative Dictionary (Blust and Trussel ongoing, see www.trussel2.com/ACD).
being cultivated on Taiwan from the initial Neolithic settlement of the island, made its dramatic appearance. The old saw that ‘absence of evidence is not evidence of absence’ could hardly have found a better application. What is perhaps more important, however, is that evidence for PAN grain agriculture was available to those archaeologists who were willing to consider the data from other academic disciplines concerned with prehistory, Peter Bellwood being preeminent among them.

**Radiocarbon dating and the phylogeny of Eastern Polynesian**

To disarm any suspicion that I might be unfairly favouring my own discipline, I will now describe a case in which the demands of radiocarbon dating in archaeology required a rethinking of a major linguistic hypothesis that had been widely accepted for almost five decades before careful reexamination of the evidence proved it unjustified.

Polynesian languages have been recognised as a discrete group since European contact, an impression that was reinforced by the relative homogeneity of culture, at least on the high islands, and Polynesian archaeologists are equally clear in recognising a distinctive cultural sequence (Kirch and Green 2001: 53–91). While the archaeology provides a fuzzier picture of the primary schism within Polynesian, since Fiji, Tonga and Samoa evidently were settled in rapid succession, the linguistics points unambiguously to Tongic (Tongan and Niue) versus the rest (Nuclear Polynesian), a division that may reflect the relative isolation of Samoa after the roughly simultaneous Polynesian settlement of the three western archipelagos. An internal family tree for Polynesian was first proposed by Elbert (1953), based mainly on lexicostatistics and this was soon followed by more detailed qualitative studies (Pawley 1966, 1967; Green 1967). The standard phylogenetic tree for Polynesian languages, which stood virtually unchallenged for nearly half a century showed Proto-Polynesian splitting into Proto-Tongic and Proto-Nuclear Polynesian (PNP), then PNP splitting into Proto-Samoic-Outlier (PSO) and Proto-Eastern Polynesian (PEP), and then PEP splitting into Rapanui and Proto-Central-Eastern (PCE), which in turn divided into Proto-Marquesic and Proto-Tahitic (Marck 2000: 2). Apart from the Tongic: Nuclear Polynesian division, perhaps the best-supported branch within this collection of languages was Eastern Polynesian, which covers an enormous expanse that essentially defines the Polynesian Triangle from Hawai’i to Rapanui to New Zealand.

Although details of this proposed phylogeny were questioned, most notably by Wilson (1985, 2012) who drew attention to evidence that linked Eastern Polynesian (EP) languages to specific Polynesian Outliers, the internal three-way division of EP remained unchallenged. As an experiment in building cross-disciplinary bridges, from August–December 2011, Terry Hunt and I co-taught a course at the University of Hawai’i titled ‘Archaeology and language in the Pacific’. An issue that continued to bother us in trying to put the archaeology and linguistics together was to find answers to the following questions: 1. Where was Proto-Eastern Polynesian spoken? 2. If Rapanui was the first language group to split off from the EP group, where was Proto-Central-Eastern Polynesian spoken? 3. How good is the evidence for Rapanui vs. Proto-Central Eastern Polynesian, and for Proto-Marquesic (the putative ancestor of Hawaiian, Marquesan and Mangarevan) versus Proto-Tahitic (the putative ancestor of Tahitian, Tuamotuan, Cook Islands Maori and New Zealand Maori, among others? Marck (2000: 3) proposed several major revisions to this model, but left the internal structure of EP intact, thus providing no answers to
these questions. What finally forced us to see this nearly 50-year-old phylogeny as flawed was the ‘sanitised’ radiocarbon chronology for eastern Polynesia proposed in Wilmshurst et al. (2011), of which Hunt was one of the contributing authors.

By restricting themselves to datable materials that provide maximum reliability, Wilmshurst et al. (2011) concluded that eastern Polynesia was settled two to five centuries later than had previously been thought. In particular, a two-phase sequence of colonisation was identified, beginning in the Society Islands AD 1025–1120 and very rapidly leading to the settlement of the rest of eastern Polynesia by AD 1190–1293. Through class discussions it became clear that if this revised chronology is valid the Pawley-Green phylogeny could not be correct; apart from the issue of identifying plausible homelands for each of the posited proto-languages, there simply was not enough time for innovations to accumulate that could distinguish Marquesic from Tahitic. Language change is inevitable, but except under the rarest circumstances it takes centuries for significant change to occur, and 100–200 years is far too short to support the scenario implied by the ‘standard’ model of Eastern Polynesian subgrouping.

Mary Walworth, a student in this class who was working in French Polynesia and who intended to write her dissertation on the language of Rapa in the Australs, took it upon herself to reopen the investigation of Eastern Polynesian subgrouping, and the harder she looked into the evidence presented for a Marquesic–Tahitic split the less convincing the argument became. She has now concluded that Eastern Polynesia was settled rapidly by speakers of a language with little dialect differentiation, and that the divergence of Rapanui from the rest is due to isolation after a nearly simultaneous settlement of the whole of eastern Polynesia. In other words, a reexamination of the linguistic evidence shows no well-supported internal subgrouping divisions, a pattern that is consistent with the whole of Eastern Polynesia minus Rapanui remaining an enormous interaction sphere through bidirectional sailing for at least a century after initial settlement (Walworth 2014). Although a similar idea was floated earlier in the archaeological literature by Kirch (1986), this is the first time an attempt has been made to defend a ‘simultaneous settlement’ view of Eastern Polynesia with evidence from both archaeology and language.

Together with the insights of Wilson (2012), it now appears likely that the Society Islands were settled directly from the northern outliers in the Solomons, and that the rest of Eastern Polynesia was peopled over a period probably not exceeding 8–10 generations, during which time contact was maintained with homeland areas and perhaps other distant locations. Following this initial period, as connections of kinship grew more distant, the interaction spheres contracted and language differentiation followed its normal course, giving rise to divergent languages and cultures. But without the prod provided by a corrected radiocarbon chronology, and its implications for language phylogeny, it is doubtful that a reevaluation of EP subgrouping would have been undertaken.

---

3 Marck states that ‘it is not clear from the archaeology where and when East Polynesia came to be settled’ (2000: 138–139). He adds that about 500 years would be needed to account for the innovations between Proto-Ellicean (ancestral to Samoan, the Ellicean Outliers and Eastern Polynesian) and Proto-Central-Eastern Polynesian, but makes no specific claim about the location of the latter. Moreover, he is explicit in claiming that ‘the divergence of Rapanui speech ... clearly preceded the differentiation of speech between the Marquesas, the Societies, the Tuamotus, the Australs (sic), and the Cooks’ (2000: 138).

4 It is only proper to add that in the years since the Wilmshurst et al. (2011) chronology was published there has been a vigorous debate of its merits within the community of Polynesian archaeologists. Much of this is summarised in Spriggs, who holds the view that the proposed settlement dates for much of Eastern Polynesia are better seen as ‘when settlement is widely established in the region’ (2014: 174) rather than as evidence of first contact. In either case, the fact that linguists responded to the revised chronology in an appropriate manner by re-examining the evidence for a Tahitic–Marquesic split and finding it weaker than earlier scholars thought is clear evidence that the pressures for interdisciplinary agreement that exist between historical linguistics and prehistoric archaeology are not unidirectional.
**PAN ‘iron’ and ‘rust’: Making sense of undeniable lexical comparisons**

The two preceding cases have shown, first, how linguistic inferences about grain agriculture preceded archaeological confirmation by decades and, second, how the demands of strictly monitored archaeological dating forced a reconsideration of the traditional internal subgrouping of Eastern Polynesian languages. Together these examples illustrate the potential benefits that each discipline can offer to the other by acting responsibly toward its own data, but simultaneously taking into account the claims of the other field.

Another insight arising from the marriage of historical linguistics and archaeology can be seen in the need to reconstruct PAN *Namat ‘iron’ and *diNaŋ ‘rust’ (Blust 2013). Words for ‘iron’ in AN languages have posed problems of interpretation at least since Dempwolff (1938) posited *besi ‘iron’, based on data from Malagasy, a few languages of western Indonesia, and a misconstrued Fijian word that no researcher today would consider cognate.

Although most previous linguistic comparisons relating to iron can now be dismissed as either flawed, or as referring to a later proto-language, the forms *Namat and *diNaŋ are unassailable. Even though reflexes of both words are restricted to Formosan languages (Taokas and Kavalan for the former, Saisiyat, Isbukun Bunun and Paiwan for the latter), these languages belong to different primary branches of the AN family, and are not now, nor do they appear to have previously been, in contact with one another. Given this distribution of apparently non-borrowed cognate linguistic forms glossed ‘iron’ and ‘rust’ in the primary sources, Blust (2013) reconstructed words with these meanings for Proto-Austronesian. Since iron-working in Taiwan developed considerably later than the initial Neolithic settlement of the island, the question arose how the meanings ‘iron’ and ‘rust’ were to be interpreted in a meaningful cultural context. An email message from Dr Hsiao-chun Hung (The Australian National University) appears to provide the answer: the earliest pottery in Taiwan was a plain red-slipped ware, and a 2005 study by Hung and the geochemist Dr Yoshiyuki Iizuka confirmed that the source of the red-slip colouring agent is hematite. Deposits of hematite are reportedly abundant in parts of Taiwan, and as the mineral form of iron oxide, its association with both iron and rust is readily understandable. What was not previously understood by linguists is that this readily decomposable mineral provided the red-slip agent for the earliest pottery in Taiwan, and that its consequent cultural importance could be reflected in linguistic forms with the generalised glosses ‘iron’ and ‘rust’, which might initially give rise to quite different culture-historical inferences.

**The history of Austronesian agriculture**

**The later invisibility of rice**

So far I have stressed that there is no single discipline that deals with human prehistory. Rather, several disciplines have developed tools for investigating the human past, including at least archaeology, population biology and historical linguistics. Some scholars in each of these disciplines appear averse to recognising the value or trustworthiness of one or more of the others. Moreover, there is a common, if often implicit, belief that the physical sciences are to be trusted more than the biological sciences, and the latter more than the social sciences. Some years ago a well-known Pacific archaeologist in Hawai‘i actually began a talk that I attended (unnoticed at first) by stating that inferences in archaeology should be prioritised over those in linguistics since archaeology deals with elements of the past that are concrete and measurable (like pottery or stone tools), while linguistics only deals with ‘words’. That remark in itself could trigger a long
debate about the nature of science and use of the hypothetico-deductive method. More recently, an email message directed to me on 27 October 2011 by a well-known population biologist contained the following jaw-dropping statement: ‘We know the methodology, the underlying assumptions, and the limitations of phylogenetic approaches. I don’t think the same can always be said to hold for other approaches to reconstructing language relationships.’

Historical linguistics has foundations predating 1820, when Jakob Grimm and Rasmus Rask demonstrated that most of the languages of northern Europe derive from the same prehistoric source community as the classical languages of the Mediterranean and India, an achievement that remains today as solid as it was when it was first proposed nearly 200 years ago, and that is sometimes ranked along with the theory of evolution by natural selection as one of the two great intellectual achievements of the nineteenth century. Needless to say, this was long before either archaeology or genetics existed as academic disciplines, and any formalised set of procedures for explaining the natural or cultural world with a history this long has repeatedly been tested, and tested severely for weaknesses and inadequacies. It therefore comes as a (somewhat annoying) surprise to historical linguists to occasionally hear from social anthropologists, archaeologists, population biologists or scholars in other disciplines that we can’t trust the evidence of language because historical linguistics is a novel discipline that has yet to prove its worth.

Given this perception, fostered in part by the insularity of much training in the sciences, it is refreshing to see scholars who recognise the value of interdisciplinary cooperation, a hallmark of much of Peter Bellwood’s work throughout his long and productive career. Despite the unquestioned benefits of interdisciplinary approaches to common problems in prehistory there are, however, risks that come from overestimating one’s understanding or command of the material of sister disciplines. An example of this inadequacy is seen in Bulbeck (2008), who argues that although Proto-Austronesian speakers on Taiwan practised cereal agriculture, they abandoned this mode of subsistence once they began to expand southward into insular Southeast Asia (ISEA), and became maritime foragers in their migrations throughout the rest of the AN world.

Bulbeck is explicit in recognising the value of a multidisciplinary approach to prehistory, but he rejects evidence that in his view is not supported by his own discipline, as seen in this passage:

I hope it is clear from my contribution that I heartily endorse a multidisciplinary approach to ISEA prehistory, and the ‘triangulation’ of the linguistic, archaeological and biological evidence relevant to the Neolithic (e.g. Sagart et al. 2005). However, I have little sympathy for taking a particular interpretation of the historical linguistics, one based on idealist culture history and the assumption of an expanding Malayo-Polynesian monoculture, and using it to overwrite the archaeological and biological evidence. In my view this reduces the number of disciplines that counts to one, which is the opposite of multidisciplinary research (2008: 48).

No explanation is given for what is meant by ‘a particular interpretation of the historical linguistics’ or ‘idealist culture history’. The 18 PAN terms relating to grain agriculture under (1) are straightforward products of application of the comparative method of linguistics. In other words, any well-trained historical linguist would reach the same conclusions based on the primary data that supports these inferences. There is nothing questionable or methodologically unsound about these forms; they are products of science as it is performed in any well-developed academic discipline.

As noted by the generalised distributional information in parentheses, all of these reconstructed forms are supported by data from at least one AN language of Taiwan, plus related languages in insular Southeast Asia. The material supporting these statements is publicly available online by visiting the AN Comparative Dictionary (ACD) (Blust and Trussell ongoing,
see www.trussel2.com/ACD). To facilitate quick reference, (2) provides a scaled-down version of the entry for PAN *pajay ‘riceplant, rice in the field’ (scaled down because the full set of data would take up far too much space in this short paper):

(2)

PAN *pajay ‘rice in the field; riceplant’

<table>
<thead>
<tr>
<th>Country</th>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taiwan</td>
<td>Atayal</td>
<td>pagay</td>
<td>unhulled rice, rice in the husk</td>
</tr>
<tr>
<td></td>
<td>Seediq</td>
<td>payay</td>
<td>riceplant</td>
</tr>
<tr>
<td></td>
<td>Saisiyat</td>
<td>pazay</td>
<td>riceplant</td>
</tr>
<tr>
<td></td>
<td>Bunun</td>
<td>paað</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Tsou</td>
<td>pai</td>
<td>riceplant, unhusked rice</td>
</tr>
<tr>
<td></td>
<td>Amis</td>
<td>panay</td>
<td>rice; the harvested grain</td>
</tr>
<tr>
<td></td>
<td>Paiwan</td>
<td>paday</td>
<td>rice (plant, grain)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philippines</td>
<td>Ilokano</td>
<td>pagay</td>
<td>rice (unhusked)</td>
</tr>
<tr>
<td></td>
<td>Kapampangan</td>
<td>pále</td>
<td>rice (growing in the field)</td>
</tr>
<tr>
<td></td>
<td>Tagalog</td>
<td>pálay</td>
<td>riceplant</td>
</tr>
<tr>
<td></td>
<td>Hanunóó</td>
<td>páray</td>
<td>rice in the field</td>
</tr>
<tr>
<td></td>
<td>Tiruray</td>
<td>farey</td>
<td>general for rice, Oryza sativa Linn.</td>
</tr>
<tr>
<td></td>
<td>Tboli</td>
<td>halay</td>
<td>unhusked rice</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borneo</td>
<td>I’a’an Begak</td>
<td>paray</td>
<td>paddy</td>
</tr>
<tr>
<td></td>
<td>Kadazan Dusun</td>
<td>paai</td>
<td>paddy</td>
</tr>
<tr>
<td></td>
<td>Kelabit</td>
<td>pade</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Kayan</td>
<td>pare</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Mukah Melanau</td>
<td>paday</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Iban</td>
<td>padi</td>
<td>riceplant, grain in husk</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnam</td>
<td>Jarai</td>
<td>pəday</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Rade</td>
<td>madie</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>Moken</td>
<td>pai</td>
<td>paddy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sumatra</td>
<td>Karo Batak</td>
<td>page</td>
<td>riceplant, rice in the field</td>
</tr>
<tr>
<td></td>
<td>Nias</td>
<td>faxe</td>
<td>riceplant, cooked rice</td>
</tr>
</tbody>
</table>
Java-Bali:
Old Javanese  pari  riceplant, paddy
Javanese  pari  riceplant, unhulled rice
Sundanese  pare  riceplant, unhulled rice
Balinese  padi  rice in the husk

Sulawesi:
Mongondow  payoy  riceplant, unhusked rice
Bare’e  pae  riceplant, unhusked rice
Tae’  pare  rice in the field
Makasarese  pare  rice in the field
Muna  pae  rice (on the stalk)

Marianas:
Chamorro  faʔi  rice (when growing)

Lesser Sundas:
Bimanesefare  fare  riceplant, rice in the field
Li’o  pare  riceplant, rice in the field
Kambera  pari  riceplant
Rotinese  hade  riceplant, unhusked rice grain
Tetun  hare  rice (plant and unhusked grain)

Moluccas:
Yamdena  fase  rice
Kamarat  hala  rice (in general)
Buruese  pala  rice (in general)
Kowiai  fasa  cooked rice, husked rice
Numfor  fas  rice (in general)

This is not the place for a mini-course in historical linguistics, but let me at least point out that a borrowing explanation for the distribution of these terms will not work because the sound correspondences, particularly those affecting *j, are both phonetically diverse and almost invariably regular, something that would be possible with loanwords only if they had been borrowed before any of these languages had differentiated from a common ancestor. Since this ancestor would have been PAN, diffusion becomes meaningless as an alternative to recognising that rice agriculture in the Tapenkeng (TPK) culture of Taiwan was continued by speakers of PMP, and then by many of its descendant communities in ISEA to the present day. Far from the empirically inadequate scenario that Bulbeck defends, what the linguistic evidence shows unambiguously is a continuation of rice (and millet) agriculture, with reflexes of PAN *pajay, *beRas and *Semay retained in the languages of the Philippines and western Indonesia, but terminological distinctions (and hence presumably importance) reduced in eastern Indonesia/West New Guinea, where sago assumes greater prominence as a staple carbohydrate, until rice finally disappears in the Pacific. The one exception to this general statement is, of course, the Marianas, which were settled directly from the central or northern Philippines by AN settlers.
who failed to transport the pig, water buffalo, and apparently chicken, but succeeded in taking rice, sugarcane, betel nut and a few other cultigens in what was probably a single one-way journey undertaken prior to 3500 BP (Blust 2000).5

Bulbeck insists that favouring the linguistic evidence at the expense of what he represents as the archaeological record violates the principle of multidisciplinary research. This may be narrowly true, but given the robustness of the linguistic evidence for PAN rice and millet agriculture and its continuation throughout ISEA to the present, it is impossible to accept the claims based on negative evidence from the archaeological record without abandoning faith in the basic principles of historical linguistics. For years archaeologists were convinced that the founding Neolithic TPK culture in Taiwan was devoid of rice agriculture. The linguistic record unambiguously contradicted this inference, and the archaeological record itself has now shown that the linguists had good reason to adopt the position they did. In short, what matters in assessing the strength of an argument is not whether the argument is based on evidence from physical science, biological science or social science, but rather how successful it is in excluding competing alternatives within its own discipline. With grain agriculture in Neolithic Taiwan there was no plausible alternative to the conclusion reached by linguists, while there obviously were other explanations for the null evidence from archaeology.

There is a general lesson in this example for how scientific method works in any field of inquiry: arguments based on negative evidence can never be conclusive because there is no logical imperative that absence of evidence provides evidence of absence. Ultimately this is equivalent to what philosophers of science have called ‘Hume’s problem’, or ‘the problem of induction’, namely that in science (unlike mathematics) it is always possible for new and unexpected observations to alter one’s previous position (Popper 1968: 28ff). Naturally, if there is no evidence of any kind for some proposition then we can say with confidence that there is no reason to believe in its validity until evidence is presented. But in the present case evidence has been presented, and the only reason that can be found for the skepticism with which it has been treated outside the community of knowledgeable historical linguists is the implicit belief that there is a hierarchy of trustworthiness in the sciences in which the physical sciences rank highest, followed by the biological sciences, and last of all by the social sciences.

Bulbeck (2008) appears not to have learned this lesson, as his interpretation of the AN diaspora is based in critical respects on the physical near-invisibility of rice south of northern Luzon prior to 2500 BP, ignoring not only the abundant linguistic evidence for continuity of grain agriculture in the AN expansion into equatorial insular Southeast Asia, but also the plain archaeological fact that the bulk of excavated sites in Island Southeast Asia to date have been in caves or rock shelters, where one would hardly expect to find agricultural remains. As an outsider to the archaeological enterprise one might reasonably ask: how many village sites prior to 3000 BP have been located and excavated in this region?

Refuting the comparative method

As already noted, linguists have good reason to dispute overinterpretations of negative evidence in archaeology that conflict with strong linguistic arguments. This is not a matter of hubris or of elevating the evidence of one academic disciple over that of another; it is a matter of collegial responsibility, and ultimately of good scientific method. Roger Blench, who is trained in African linguistics, is sufficiently aware of this issue to recognise that anyone who wishes to support

---

5 As noted in Blust (2000: 106–107) Chamorro has a native word reflecting PAN *baRiuS ‘typhoon’. Since the typhoon zone extends from the central Philippines to southern Japan, with rare strikes in Mindanao and virtually none further south, and since Chamorro is clearly a Malayo-Polynesian (non-Formosan) language, a process of elimination leads us to a source in the central or northern Philippines.
Bulbeck’s maritime foraging model must reconcile the negative arguments from archaeology with the positive arguments from linguistics—arguments that Bulbeck himself simply did not consider. In Blench (2012) he does this through a tour de force of unrestrained assertion: among other things, we are told 1) ‘That the Austronesians, far from being agriculturalists, were fisher-foragers’, 2) ‘That the dogs, pigs and chickens supposedly characteristic of this expansion reached Island Southeast Asia through alternative routes’, 3) ‘That the linguistic reconstructions that appeared to support this model are in fact mosaics of loanwords’, and 4) ‘That the spread of Austronesian languages was due to a powerful religious/lifestyle ideology which assimilated indigenous speech communities and that this can be detected from material culture’ (Blench 2012: 128).

Let me begin with points 2) and 3), since these will go a long way toward addressing point 1). Starting with the claim that the ‘Pacific Clade’ of pigs originated in ‘Laos, Yunnan and far Northwest Vietnam’ (Larson et al. 2010 and earlier work), Blench (2012: 134) concludes that AN speakers moving out of Taiwan ‘would have adopted dogs, pigs and chickens from MSEA [Mainland Southeast Asia] and then carried them eastwards to Melanesia and Oceania’.

 Entirely apart from the odd distinction between ‘Melanesia’ and ‘Oceania’ in this quote, which hardly instills confidence in his knowledge of the area, Blench demonstrates a remarkable ability to sweep away large quantities of conflicting linguistic data by simply declaring without evidence or argument that it is all due to borrowing. The Austronesian Comparative Dictionary (Blust and Trussell ongoing, see www.trussel2.com/ACD) is an open-access site that continues to grow on a weekly basis, but is already the largest research project ever undertaken on the AN languages (Blust and Trussell 2013). The entries for ‘dog’, ‘pig’ and ‘chicken’ in the ACD are: PAN *asu ‘dog’, PAN *babuy ‘pig’, PAN *beRek ‘domesticated pig’, and PAN *manuk ‘chicken’. The first of these is currently documented by 78 forms and the second by 74 forms in attested languages, in both cases reaching from Taiwan to the Moluccas. The third reconstruction (*beRek) is documented by 26 terms reaching from Taiwan to the central Solomon Islands, and *manuk by 136 terms reaching from Taiwan to Hawai’i. Scholars who use the ACD with care know that loanwords are carefully distinguished from native forms. This is done in two distinct ways. First, apart from the main file (‘Cognate sets’), which includes reconstructions and supporting evidence for over 6,600 base forms and more than 21,000 total reconstructions (bases plus affixed words) there are several separate files, including one for ‘noise’ (resemblances due to chance), and another for loanwords that are sufficiently widespread to introduce the danger of false reconstructions based on loan distributions, whether these be due to the diffusion of lexical innovations from particular AN languages (such as Malay), or borrowings from outside the AN language family (as widespread terms for ‘papaya’, ‘guava’, and ‘pineapple’). In addition, many entries in the dictionary contain notes, and with very few exceptions forms that show irregularities in the expected sound correspondences (usually because of borrowing between related languages) are cited in the note separately from the main entry, leaving all or virtually all items in the main entry free from any problems with regard to the regularity of the sound correspondences.

Despite the care with which this material has been treated, Blench, who nowhere cites the ACD, notes that ‘dogs are again conspicuous by their absence in the early archaeology of ISEA, except for Timor … Could all occurrences of the *asu root represent a semantic shift or borrowing?’ (2012: 137). After acknowledging that this seems unlikely, he observes that the proto-form for ‘dog’ in Tai-Kadai appears to be a loanword from Hmong-Mien, and then concludes abruptly that if this could happen in Tai-Kadai ‘this term may have been absent in PMP and all those occurrences of #asu are in fact loanwords reflecting early contact with Austroasiatic speakers (and the subsequent spread of the term once borrowed). Although this contradicts established wisdom in AN scholarship, it shows greater congruence with the archaeological record’
To this assertion, Blench then appends the following footnote, evidently intended as a metaphorical nail in some imagined coffin: ‘# is a linguistic convention used here to indicate a quasi-reconstruction, a form derived from rapid inspection of cognates’ (2012: 144).

At no point does Blench identify a potential Austroasiatic source for PAN *asu, nor does he mention that the word shows regular phonological developments not just in the 78 languages from which data is cited in the ACD, but from several hundred others, thus providing no empirical basis for his claim that the word has been borrowed from one AN language into another. In short, it is not ‘established wisdom in AN scholarship’ that he proposes to challenge, but rather the massive evidence that *asu must be reconstructed to explain the distributional data in daughter languages. The offhand remark that this reconstruction has been ‘derived from rapid inspection of cognates’, which clearly implies slipshod scholarship, is a glaring misrepresentation of the meticulous work of Otto Dempwolff (1934–1938), who laid the foundations of Austronesian comparative linguistics, and of the generations of scholars who have followed in his footsteps over the past eight decades.

It is apparent from Blench’s ‘wave of the hand’ dismissal of fundamental products of the comparative method that his agenda is to reconcile the linguistic record with the negative evidence of archaeology by pursuing an ‘absence of evidence is evidence of absence’ argument, and simply sweeping the massive countervailing linguistic evidence out of sight. But, as with Bulbeck’s earlier argument, Blench’s use of negative evidence to rewrite AN culture history is a castle built on sand. As has already been shown with TPK in Taiwan, there simply is no logical imperative that absence of evidence is evidence of absence, and for a linguist, albeit one who is not a specialist in the AN language family, to dismiss valid linguistic evidence in order to reconcile it with zero visibility in the archaeological record is surprising, to say the least.

Blench’s unfamiliarity with the comparative method is, if anything, even more evident in his treatment of *babuy and *beRek. The reader is told (Blench 2012: 137) that ‘[a]s with “dog” there is a widespread term in Austroasiatic (AA) which is apparently cognate either *C-liik or *C-lek’. The casual reader, particularly the non-linguist, probably will fail to notice that Blench has not, in fact, shown any AA form that resembles PAN *asu, let alone one that can be demonstrated by recurrent sound correspondences (hence by the comparative method) to derive from a common proto-language. With regard to ‘pig’, Blench also neglects to mention that the ‘reconstruction’ he cites is his own impromptu speculation, not the product of knowledgeable scholarship by AA specialists who have identified sound correspondences and used them to propose a PAA reconstruction, let alone a form ancestral to both AA and AN languages. As with ‘dog’, both terms for ‘pig’ exhibit regular sound correspondences through many widely separated languages, showing that with very few exceptions they are directly inherited. This is particularly true of *beRek, which has distinctive reflexes of the vowels and *R that could easily give away the status of loanwords (and sometimes does), yet it is regular in its development through scores of languages reaching from Taiwan (Tsou fraʔa, Puyuma tarak ‘domesticated pig’), to the Philippines (Isneg hagoʔ ‘young hog’, Pangasinan belék, Kapampangan abyák ‘suckling pig’), to Borneo (Paitan vogok, Kelabit tarak ‘domesticated pig’) to Old Javanese as recorded in palm leaf manuscripts.

6 Referring to the late Harry L. Shorto, Paul Sidwell (email of 4 November 2014 informs me that ‘Shorto proposed *lik and “cur “pig” and “cetʔ “boar”’. *lik is pretty solid, it is in 5 branches. “cur is restricted to Eastern languages, mainly Bahnaric, and looks like a local word. “cetʔ “boar” is in Adlian and Bahnaric. None are in Munda, apparently. I suspect that “lik is imitative’. Note that the term ‘cognate’ is used here with reference to both ‘dog’ and ‘pig’, thus assuming that the Austro hypothesis is valid, and that specific Proto-Austric lexical forms can be posited, whereas for ‘dog’ Blench first claimed that PAN *asu is a loanword from an unidentified AA source, not a retention from Proto-Austric. These are not the only inconsistencies that will trouble linguists who read this paper. With regard to possible AA cognates of PAN *beRek, for example, Blench holds that ‘The fricatives in Taiwanese may well be cognate with forms such as Peer struk’ (Blench 2012: 139). The fact that ‘Taiwanese’ is a form of Minnan Chinese, not an alternative name for the Formosan aboriginal languages, seems to escape him, but more seriously, cognition refers to morphemes, not to phonemes – an elementary observation that can easily be used to distinguish someone who understands historical linguistics from someone who does not.
dating from the beginning of the ninth to the end of the fourteenth centuries (wök ‘pig, hog, boar’), to a range of Oceanic languages, as Tigak (New Ireland) vogo, Woge (New Guinea) boro, and Nggela (central Solomons) mbola, all glossed simply as ‘pig’. Blust (1976) noted that reflexes of PAN *babuy and *beRek have different distributions that may well correlate with their distinct meanings: AN speakers introduced domesticated pigs and the name for them into the Pacific, whereas the word *babuy, which often refers to wild pigs, was lost after leaving the Moluccas, and only later reacquired in Chamorro, Yapese, and Palauan, where phonological irregularities show that the words in these three languages are loans.

Whatever merit the genetic studies of Larson et al. (2010) and others may have, they cannot be used to deny that AN speakers had terms for both domesticated and wild pigs during the initial Neolithic settlement of Taiwan and that these terms persisted during the southward migrations into insular Southeast Asia – the evidence for this is simply too strong to dismiss out of hand, and it is now up to the scientific community to try to understand why the evidence of DNA and the evidence of language appear to be at loggerheads in relation to this issue.

With regard to reconstructions relating to domesticated plants such as banana, sugarcane and taro, Blench concludes: ‘When it was supposed that Austronesian was a gradual demographic expansion, it was inevitable that these subsistence terms were seen as “true” reconstructions; with our current perspective on dispersal they must now be seen as interesting fakes’ (2012: 142). Again, to call such solidly supported linguistic reconstructions as PAN *beNbeN, PMP *punti ‘banana’, PAN *CebuS ‘sugarcane’, PMP *tales ‘taro: Colocasia esculenta’ ‘fakes’ shows little understanding of the sound correspondences, and should serve to indicate that Blench’s approach to the floral evidence is identical to that for the faunal evidence: wherever the linguistic data conflicts (or appears to conflict) with the data from either archaeology or genetics the linguistics must be wrong, and decades of careful analysis and discussion are jettisoned as products of misguided scholarship.

To cap this misadventure in freewheeling imagination, Blench argues that the Austronesian diaspora was driven by ‘a powerful religious/lifestyle ideology which assimilated indigenous speech communities and that this can be detected from material culture’ (2012: 128, 134ff).

The interested reader is hopeful that some truly convincing argument will be presented for this novel claim. In fact, all that is offered is a pedestrian reference to the ubiquity of the well-known squatting figures in traditional iconography. Blench calls these bulul, without stating the source of the term, but it appears to be from Ifugaw bulul ‘wooden statue in which one of the rice culture Deities of the Underworld is believed to reside in order to guard the crops’ (Lambrecht 1978: 95), bulul ‘a rice-god idol, carved of wood and with a resident spirit ... These idols are kept in a rice granary by the rich as a guardian to assure an ample supply of rice from the store. They are brought out during harvest and offered fat and blood from a sacrificed pig’ (Newell 1993). Since this iconographic motif reportedly extends from the northern Philippines to Polynesia and is said to have no distributional parallels in the art of Africa or other regions, Blench opines that its spread must have been comparable to that of the iconography in world religions such as Christianity, and is thus due to a similar proselytising zeal. The only other ‘arguments’ offered for this conclusion are the assertions that the Pama-Nyungan expansion in Aboriginal Australia was ‘associated with a new type of stone blade technology, but most importantly with a type of clan organisation and with a pattern of singing’ (Blench 2012: 136), and that the Arawakan migrations from the eastern edge of the Andes (montaña) to the Antilles were associated with a particular style of petroglyphs in the Amazon basin. In short, what is given as support for this
claim is a few lines of random observations loosely connected by something resembling free-association more than scientific argumentation. Ironically, Blench seems happily unaware that the Ifugaw bulul or bulul, the eponymous inspiration of what we might call his ‘Bulul religion’, is the material residence of a rice culture deity?"'

Conclusions

In conclusion, with the growth of the ACD and the appearance of The Lexicon of Proto Oceanic, it is clearer now than ever that archaeology and historical linguistics in the Austronesian world need one another to reach beyond the limitations of either discipline as an independent window on the human past. As shown by the work of Bellwood, and others (e.g. Atkinson and Gray 2005; Gray et al. 2009, 2010), much can be gained from interdisciplinary cooperation between scholars in archaeology, population genetics, evolutionary biology, and historical linguistics.

At the same time caution is in order. While I feel no reservations about challenging the claims of a sister discipline if it conflicts with strong positive evidence from my own field, I would hesitate to assume the role of arbiter between competing hypotheses in archaeology or population genetics if I did not have reason to do so based on linguistic data that I control with a high level of confidence. Yet this kind of overstepping of disciplinary bounds has happened in some cases, as with Spriggs (2011: 511–514), who declared decisively, largely on the basis of a single paper (Donohue and Grimes 2008) that the Central Eastern Malayo-Polynesian node first proposed for the Austronesian family tree in Blust (1982) has been abandoned by the field, and that the AN languages of eastern Indonesia consequently constitute many primary branches of Malayo-Polynesian, thus dramatically resetting the stage with regard to implications for archaeology. Yet there are serious problems with the Donohue and Grimes proposal that were exposed promptly (Blust 2009; not noted by Spriggs). A further attack on CEMP by Schapper (2011) was similarly answered in short order (Blust 2012). The reality is that the CEMP node has not been dismissed by knowledgeable comparativists working in the area, and one can only wonder what triggered this imaginary ‘revolution’.9

All of this goes to show that historical linguists and archaeologists have what can be called an ‘academic alliance’, but owing to fundamental differences in the observational basis of the two fields it is an uneasy alliance, and one that is likely to remain this way well into the future.

References


7 Chen (1968: 378–405), who is not cited by Blench, devotes a fairly detailed discussion to ‘Figures in squatting posture’, along with numerous line drawings. Far from having associations with an aggressively proselytising religion, he argues that these figures (which also occur in the material culture of Shang and Zhou dynasty China, and on the northwest coast of North America) are naturalistic portrayals of the sitting posture that was customary before the common use of stools or chairs, namely squatting on the ground.

8 In fairness, it should be noted that Spriggs (2012) has redressed several of these shortcomings.

9 Cf. Ross, who states ‘Donohue & Grimes (2008) critique Blust’s PCEMP innovations and suggest that the PCEMP node is a chimera that should be done away with, with the consequence that Blust’s Central MP languages would simply be treated as further western MP subgroups. Blust (2009) makes a strong response, and although some CEMP lexical innovations are not particularly convincing, first-syllable coda-loss in reduplicated monosyllables, the marsupial terms, and the remaining lexical innovations do, it seems to me, pass muster’ (forthcoming, fn. 6).


