Agriculture and phylic dispersals; re-evaluating the evidence

Presented at the meeting

Us and Them: Modelling past genetic, linguistic, and cultural boundaries

Bordeaux 15–17, May 2008

and subsequently revised

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

This printout: May 23, 2008
# TABLE OF CONTENTS

1. INTRODUCTION .................................................................................................................. 1
   1.1 Claims about agricultural expansions .......................................................... 1
   1.2 Is it necessary for there to be any engine of language phylum expansion? ... 1
   1.3 Demographic expansion versus cultural transmission ............................... 2

2. THE REIFICATION OF LANGUAGE PHYLA .................................................................. 2
   2.1 How far is ‘Austronesian’ now a linguistic concept? ................................ 2
   2.2 When scholars disagree .............................................................................. 3

3. METHODOLOGICAL GAPS IN THE ARGUMENT ..................................................... 3

4. REVIEW BY LANGUAGE PHYLM ............................................................................... 5
   4.1 Africa ............................................. 5
      4.1.1 Khoesan ........................................ 5
      4.1.2 Niger-Congo ................................. 5
      4.1.3 Nilo-Saharan ................................ 7
      4.1.4 Afroasiatic ................................. 8
   4.2 Eurasia .......................................... 9
      4.2.1 Dravidian ....................................... 9
      4.2.2 Indo-European ............................ 9
      4.2.3 Uralic ......................................... 9
      4.2.4 Altaic [?] ...................................... 10
      4.2.5 Kartvelian .................................... 10
      4.2.6 North Caucasian .......................... 11
      4.2.7 Sino-Tibetan [=Tibeto-Burman] ...... 11
      4.2.8 Andamanese ............................... 11
      4.2.9 Austroasiatic ............................... 12
      4.2.10 Daic [=Tai-Kadai, Kradai] ............ 12
      4.2.11 Miao-Yao ................................. 13
   4.3 Pacific ........................................ 13
      4.3.1 Austronesian .............................. 13
      4.3.2 Papuan ....................................... 14
      4.3.3 Trans New Guinea ....................... 14
      4.3.4 Australian ................................. 14
      4.3.5 Pama-Nyungan ......................... 15
   4.4 New World .................................. 15
      4.4.1 Eskimo-Aleut .............................. 15
      4.4.2 Na-Dene ..................................... 15
      4.4.3 Mayan ....................................... 16
      4.4.4 Otomanguean ............................. 16
      4.4.5 Mixe-Zoque ............................... 16
      4.4.6 Uto-Aztecan .............................. 16
      4.4.7 Cariban ..................................... 16
      4.4.8 Arawakan [=Maipuran] ............... 17

5. SYNTHESIS ................................................................................................................. 17
   5.1 The pattern of isolates and small families ............................................. 17
   5.2 Phyla: dates, homelands, motives for expansion ............................... 19

6. CONCLUSIONS ........................................................................................................... 20

REFERENCES .................................................................................................................. 23
TABLES

Table 1. Proto-Kartvelian terms relating to agriculture .......................................................... 10
Table 2. Proto-Otomanguean subsistence reconstructions ...................................................... 16
Table 3. Isolates, small phyla by continent ............................................................................ 18
Table 4. Language phyla and groupings of the world .............................................................. 18
Table 5. Language phyla of the world and linguistic evidence for subsistence ..................... 19
Table 6. Language phyla of the world, dates, homelands etc. ................................................. 20

FIGURES

Figure 1. Niger-Congo tree ....................................................................................................... 6
Figure 2. Internal structure of Afroasiatic ................................................................................ 8
Figure 3. Classification of the Indo-European languages ......................................................... 9
Figure 4. Internal classification of Daic .................................................................................. 12
Figure 5. High-level structure of Austronesian according to Blust (1999) .............................. 13
Figure 6. Simplified tree of Australian languages ................................................................. 15
Figure 7. The Cariban languages ............................................................................................ 17
Figure 8. Foraging, the transition to farming and language diversification ......................... 22
1. Introduction

1.1 Claims about agricultural expansions

Associated principally with archaeology are a number of highly questionable claims concerning the antiquity of language phyla and the motivation for their expansion. This idea has a long history within Indo-European, but has most recently been associated with the work of Peter Bellwood who has energetically propagated the notion that many language families expanded as a result of both demic diffusion and the spread of farming (e.g. Bellwood 2005, Bellwood & Renfrew 2002). While such explanations might be seriously entertained in a small number of specific cases, the case for attributing this hypothesis a more global explanatory power is questionable. It rarely addresses actual linguistic evidence and indeed, in many cases, the data appears to contradict the model. At the same time, linguists do not always present their results in a most accessible way and sometimes offer reconstructions that are chronologically improbable or culturally unrealistic. Linguists also disagree, rather forcefully in some cases; the case of Sino-Tibetan springs to mind. No wonder many archaeologists and linguists on the margins of the debate simply shrug their shoulders.

Two questions spring to mind; is it in fact the case that any unitary explanation is likely to be productive and what does the linguistic evidence actually suggest in the case of phyla already analysed? As it happens, both can be answered empirically, by reviewing the literature for individual language phyla. This is not as easy as it sounds, both because the literature is scattered in many languages, but also because of major disagreements about classification and reconstruction. This paper is an attempt to begin this task to produce a more nuanced account of the expansion of language phyla.

One motivation for attempting this task is to counter the capture of the agenda for this type of global history. As a byproduct of their argument, archaeologists may perpetuate old and broadly discredited hypotheses. For example, the world map of language phyla accompanying Ruhlen (1991) essentially represents the ideas of the linguist Joseph Greenberg (1915-2001) and it has been recopied in archaeological texts (e.g. Bellwood & Renfrew 2002) despite the fact that the classification it represents has hardly been endorsed by any subsequent linguists. At the other end of the spectrum lies the Ethnologue (http://www.ethnologue.com/web.asp), a remarkable compilation maintained by the Summer Institute of Linguistics, which is a relatively up-to-date catalogue of actual languages, but replicates an extremely idiosyncratic approach to classification.

1.2 Is it necessary for there to be any engine of language phylum expansion?

It might be asked whether it is necessary to adduce any motive for language phylum expansion other than natural population growth. The answer is that in most parts of the world, the language families are manifestly recent. Indo-European, Afro-Asiatic, Austroasiatic, Austronesian, Kartvelian, Mixe-Zoque and Mayan are all phyla that are relatively tightly knit and have clearly expanded recently, eliminating a great deal of prior diversity. Indeed this is one of the factors that makes a link with the rise of farming credible; what would be behind this if not agriculture? If agriculture is ruled out then it is reasonable for opponents of the hypothesis to present an alternative; close-knit language phyla do not exist simply by chance. Other alternatives are superior military organisation linked with expansionist ideologies, new technology (metals, maritime techniques, bow and arrow even lithics), climate or environmental change leading to new resource availability, changes in health status and religious and social ideas. Some of these are more easily tested against archaeological and linguistic data than others.

However, the question is not without a logic. In the case of Papuan, non-Khwe Khoesan or Uralic, the absence of a unitary phylum may indeed reflect nothing more than slow demographic growth among foragers with a subsequent introduction of agriculture and herding (in the case of Papuan and Uralic). These

---

1 This paper was first presented at the meeting Us and Them: Modelling past genetic, linguistic, and cultural boundaries Bordeaux 15–17, May 2008 and subsequently revised. I would like to thank Francesco D’Errico for inviting me to the meeting and the audience for their comments. In particular, discussions with Love Eriksen have helped me update my presentation of Arawakan

2 For example, ‘Australian’ is listed a phylum with the same status as Indo-European, a view espoused by almost no-one in the field, though shown that way on Ruhlen’s map.
may be no need to adduce further reasons to explain their geographical and linguistic situation at the macro-
level.

### 1.3 Demographic expansion versus cultural transmission

Associated with the language/farming dispersal hypothesis is the idea of an association with demographic
expansion, ‘demic diffusion’ in the language of its advocates. In principle, this is realistic; we know that in
the early stages of the evolution of farming, farmers move regularly, sometimes in large circular patterns, to
allow for the regrowth of fallow, but often pioneering new areas of untilled land. Similarly, pastoralists or
fishermen may depend on a resource in a particular area for some time, but the year it fails they explore new
territory and their movement patterns and dispersal often change irrevocably. Nonetheless, it is clearly also
true that language can spread independently of mass migration; the vast majority of Americans have not
absorbed American culture by direct transmission from English-speaking forebears, although they also
migrated. Similarly, many Chinese populations today were clearly formerly speakers of languages of other
phyla and have gradually ‘become’ Chinese over the millennia.

This is not to question the reality of demographic expansions; they clearly occur. However, the relationship
with a subsistence pattern has to be demonstrated on a case by case basis. This is not difficult when the
populations and their languages expand into territory only occupied by foragers or is uninhabited. There is
no interpretation of the early phases of the Bantu expansion in Central Africa that suggests they were in
competition with in situ farmers. Settlement sites appear in former forager territory. Similarly, the
Polynesians sailed to previously uninhabited islands and the link between farming, demographic expansion
and language is hardly in doubt. However, even further up the Austronesian tree the situation is not so
simple. The claim is that after the settlement of Taiwan, maritime Austronesian-speakers moved southward
to the northern tip of Luzon, gradually expanding their agricultural settlements until they had settled the
entire archipelago, while simultaneously exploring other landfalls by sea. The justification for this is the
appearance of new types of pottery (red slip ware). But opponents argue that this is not evidence for
migration but for the gradual evolution of a complex trade network and that it was ideas and artefacts that
were spreading, not people.

Clearly this is not an issue to be resolved unambiguously. When people move, artefacts and ideas also
migrate. Proving demographic expansion would require the sort of large scale quantitative analysis and
sampling that is unlikely to be available for most parts of the world. However, most reasonable onlookers
would accept clusters of villages moving and spreading with characteristic new types of material culture as a
priori evidence for this type of migration.

In recent years, a quite different type of evidence has been adduced in support of migrationist models, that of
human genetics. Both paternal and maternal DNA can be assessed for comparable lineages and thus for
evidence for genetic connections between geographically dispersed populations and thereby migration. This
type of study provides evidence that the Austronesian expansion was a real migration, both from Taiwan to
island SE Asia and on to Madagascar. Studies on the island of Bali show that the Hindu culture was brought
by a physical movement of populations from India, and did not develop from cultural transmission, via trade
networks or other (Lansing et al., in press). Interesting cases exist where paternal DNA shows distinctive
patterns, for example in China, where its presence in southern populations points to a preferential migration
of males southwards and presumably intermarriage with resident non-Chinese populations (Blench, in
press). Not all proposed migrations are so easily supported, but the evidence of DNA can provide additional
evidence to underpin or negate hypotheses of demographic expansion.

### 2. The reification of language phyla

#### 2.1 How far is ‘Austronesian’ now a linguistic concept?

A process almost unnoticed in this type of argument is the gradual reification of language phyla. Concepts
such as ‘Austronesian’ or ‘Bantu’ begin as purely linguistic hypotheses. Linguists notice the relationships
between languages and these gradually develop into a named grouping, often over centuries. Ottar of Heligoland first recorded language similarities that would lead to the recognition of Uralic in the ninth century. The first evidence for Bantu was put forward in the seventeenth century, but it was not until the mid-nineteenth that linguists accepted the existence of a widespread language family. Relandus first pointed to the links that would begin Austronesian, but it came into existence with Dempwolff in 1938.

Acceptance by linguists is followed by the genesis of a hypothetical people, Indo-Europeans, Polynesians, Bantu or Semites. As archaeologists and other interested in prehistory get drawn in, so such ancestral peoples acquire characteristics, based partly on linguistic reconstructions but increasingly on non-linguistic elements. Interdisciplinary books are published drawing on a variety of disciplines and these peoples are called into existence. Such constructs lie behind the farming dispersals hypothesis; speakers of a developing language phylum have a society, an associated set of material culture and can be visualised as migrating or undertaking pioneer farming.

2.2 When scholars disagree

This is a persuasive vision and it is sometimes difficult to remember that these people who appear on the covers of illustrated books sailing outriggers or thundering across the plains of Central Asia are only linguistic constructs and sometimes shaky ones at that. Much has been said about the culture of Altaic or Uralic speakers, but both of these linguistic phyla have been questioned by responsible scholars. For example, Janhunen (1994, 2003) has questioned the very existence of Altaic, whereas Starostin et al (2003) have published thousands of reconstructions of Macro-Altaic (including Japanese) and have confidently attributed agriculture to proto-Altaic. Competing reconstructions of Nilo-Saharan have drawn very different pictures of the culture of speakers of the proto-language (Bender 1996; Ehret 2001). Ehret (2001) claims that agriculture can be reconstructed for proto-Nilo-Saharan, a finding supported by no other researchers (see review in Blench 2002). Similarly, Militarev (2002) has claimed that Afroasiatic is an agricultural expansion and supported this assertion with a list of proto-forms not identified by other scholars. Apart from the difficulties of making sense of this in archaeological terms, the reconstructions themselves may be quite incredible, depending on semantic equivalences that are ad hoc and constructed for each individual etymology. It is essential that the proto-cultures nascent in the imagination of researchers are rooted in actual linguistic facts, rather than a house of cards of supposition held up by a scatter of dubious data.

3. Methodological gaps in the argument

What preconditions are required for there to be a reasonable a priori link between the expansion of a language phylum and agriculture (or indeed any other subsistence system, such as fisheries or pastoralism)? The answers may seem obvious, most but published models do not clearly adhere to them. They are;

1. That there be an incontrovertible phylum
2. That the phylum be sufficiently large for useful conclusions to be drawn from reconstruction
3. That the internal structure of the phylum is generally accepted and from this some assessment of the homeland and general direction of migration is available
4. For reconstructions to exist for a significant number of items including those of an ecological nature that broadly support the outline in 3.
5. That reconstructions exist of the principal crops, trees, livestock species or other subsistence items relevant to the hypothesis
6. That reconstructions exist of items suggestive of farming rather than just gathering wild relatives of the crops
7. That well-dated archaeobotanical materials exist that correspond to the reconstructions in areas roughly coincident with the proposed homeland
8. That no other competing hypothesis be available to explain the data equally well

1. Broadly speaking, language phyla can be divided into those which are almost universally accepted, those which are debatable and those which are not phyla at all but merely geographical groupings. Examples of
those universally accepted are Indo-European, Kartvelian, Austronesian, Austroasiatic, Daic, Miao-Yao, Niger-Congo, Afro-Asiatic, Mayan, Na-Dene, Uto-Aztecan, Arawakan, Dravidian, Pama-Nyungan and Sino-Tibetan. In the debated category are Penutian, Nilo-Saharan, Altaic, Macro-Khoesan (i.e. Khoesan including Hadza and Sandawe) and Trans-New Guinea. ‘Geographical’ groupings are Andamanese, Australian and Papuan which are not commonly seen as having genetic unity. In addition, there are macro-phyla, bundlings of multiple phyla, most notably Nostratic (Eurasiatic) and Amerind. These are not widely accepted by the linguistic community and, even if real, would have a time-depth too great to admit significant cultural reconstruction.

2. Not all language phyla are large; Eskimo-Aleut, Kartvelian, Koryak-Kamchadal and many groupings in the New World have just two or three members and thus cannot be reconstructed to any great time-depth. Even if agriculture, herding or fishing can be reconstructed for their proto-language this does not carry much information as we know these methods of subsistence are older than any hypothetical date emerging from the reconstruction process.

3. The internal structure of most language phyla is debated; and in some disagreement is so serious that uncontroversial reconstruction is impossible. Sino-Tibetan, apparently one of the oldest and most well-recognised phyla, is either treated agnostically or disputed. Matisoff (2003) puts forward a ‘geographical’ model of Sino-Tibetan after representing an initial branching with Sinitic languages. Van Driem (2001) is even more agnostic, mapping the ‘fallen leaves’ of Sino-Tibetan (Tibeto-Burman in his terminology) with numerous groupings and no internal structure. Other disputed phylic structures include Indo-European, Afroasiatic, Nilo-Saharan and Arawakan. Still others, such as Austroasiatic and Niger-Congo, have published trees but no justification. The significance of this is that there can be no convincing reconstructions of a proto-language without a hierarchical structure. Linguists can extract common forms (i.e. likely abstractions based on synchronic attestations) but these are not suitable for the reconstruction of prehistory.

4. If there is a dataset of proposed reconstructions then items of significance for ecology and subsistence should be a small proportion of a large dataset which will mostly include more common lexical items. Regular reconstructions of such items increase confidence in more specialised lexemes.

5. Again, this point may seem obvious, but in many cases, the actual data is highly defective. For example, the reconstruction of agriculture in proto-Bantu, an apparently obvious case of a Neolithic expansion, is hardly overwhelming. There are no clear cases of cereals or tubers and some pulses that have reconstructible forms are possibly wild rather than domestic. Afroasiatic has a clear case of a reconstruction for ‘cow, cattle’ but no other livestock and certainly no cereals or other crops.

6. To demonstrate that a phylum or subgroup is associated with true farming as opposed to foraging it is not enough to reconstruct crop names. Where crops are domesticated from indigenous species, then the name often is transferred from the gathered wild plant to the cultigen with no evident linguistic discontinuity. For example, it is claimed that ‘wheat’ can be reconstructed in proto-Indo-European, but wild wheat can be gathered almost throughout its range of proposed homelands.

To those familiar with the data on world language phyla, it will be immediately apparent that no language phylum entirely meets these criteria. Probably Austronesian is closest to these prerequisites; no wonder then it is most commonly cited in developing this type of argument.

3 ‘Universally accepted’ refers to scholars of the phylum. It can happen that researchers with reputations in quite different language phyla cast doubt on the existence of a phylum. For example, Dixon (1997), a researcher on Australian languages, states that Niger-Congo is ‘a bundle of overlapping isoglosses’, but there is no reason to take such a judgment seriously.
4. Review by language phylum

4.1 Africa

4.1.1 Khoesan

The Khoesan [= Khoisan] languages are easily the most problematic phylum in Africa, primarily because it is almost certain that they do not constitute a genetic grouping. The languages themselves are spoken by small scattered populations in south-western Africa and they are under threat from their dominant neighbours. We know of the existence of some 100 Khoesan languages from records, yet only 30 are spoken today (Güldemann & Voßen 2000:99).

The only part of Khoesan that is incontrovertible is Central Khoesan or Khwe, reconstructed by Voßen (1996). The Khwe remained foragers until European contact, but some impulse caused the expansion of a related group of languages, probably quite recently, as little as >2000 years ago. This almost certainly eliminated a rather more diverse set of languages spoken in the Central area, relegating diverse languages of branches to the periphery (Eastern ‡Hôã, Kwadi, !Xoo). It is possible to reconstruct domestic animals in Khwe and it is more than possible that that its expansion was the result of contact with pastoral populations reaching this part of Africa, as it is at about this period that cattle, sheep and pottery appear in the archaeological record (Blench 2006, in press). Guldemann (2004) has argued that Kwadi forms a subgroup with Khwe but this was apparently at a period so remote that contact with pastoralism had not yet occurred.

4.1.2 Niger-Congo

The Niger-Congo phylum is the largest in the world, including over 1500 languages. In geographical spread and internal diversity it can be compared with Austronesian (Williamson & Blench 2000). However, unlike Austronesian, it remains very poorly known, especially in West Africa, and many crucial and difficult to classify languages are represented by short wordlists rather than dictionaries. Although there are two major comparative sources, Westermann (1927) and Mukarovský (1976-77), neither worked with the modern concept of the phylum and exclude many branches that are now considered integral to Niger-Congo. Although Greenberg (1963) is usually credited with the modern concept of Niger-Congo, he followed Westermann in most ways, continuing to propagate misleading classificatory errors.

A crucial error that emerges from Greenberg’s picture of Niger-Congo is the notion that the proto-language must have had noun-classes and the related assumption that any language recalcitrant enough not to have such languages today clearly mislaid them in the past. Although this is clearly the case for some language branches, such as Kwa, elsewhere in Dogon, Mande and Ijo there is absolutely no trace of nominal affixing. Figure 1 shows a new version of the Ị iger-Congo phylogenetic tree, based on Blench (2006).
As with other African phyla, it is relevant to note the late inception of agriculture in sub-Saharan Africa and the impossibility that the early stages of Niger-Congo were in any way agricultural (Kahlheber & Neumann...
However, it is possible to reconstruct ‘bow’, ‘arrow’ and ‘spear’, which makes it likely that the early stages of Niger-Congo were foraging. Blench (2006) argues that the expansion of Niger-Congo should be related to the introduction of microlithic technology and the improving climate at the beginning of the Holocene some 11,000 years ago.

For unclear reasons, archaeological sources often use the outdated name ‘Iger-Kordofanian’ and focus on Bantu (e.g. Renfrew 2002). However, Bantu is a remote subgroup of Niger-Congo and very late in its evolution. Even if it were the case the Bantu was indeed an agricultural expansion, this would be like choosing Germanic to represent the whole of Indo-European. Indeed the demic diffusion hypothesis is indefeasible if it is permissible to pick and choose the subgroup convenient to the argument.

Even the case that Bantu was initially an agricultural expansion can be open to doubt. The reconstruction of crop names in proto-Bantu is uncertain (Blench 1996) and many of those that appear to reconstruct have wild relatives in the region. The first pottery found in the Equatorial forest is in low-density scatters along rivers and not in association with expanding farming populations. No archaeological sites in the forest have yielded archaeobotanical remains unequivocally associated with agriculture at the sort of date usually associated with the Bantu expansion (Eggert et al. 2006). Interestingly, the situation is quite different with fish names. Mouguiama-Daouda (2005) has recently undertaken detailed reconstructions of fish-names in proto-Bantu and produced convincing evidence for a number of species. Without concomitant reconstructions for traps and other fishing gear this is hardly proof of an aquatic expansion, but it is certainly suggestive. It may well be that the first phase of the Bantu expansion was only agricultural with fishing and intensive gathering the principal subsistence strategies. Some centuries later, with the introduction of SE Asian staples such as the plantain and the cocoyam, agricultural expansion then became established.

### 4.1.3 Nilo-Saharan

Nilo-Saharan is the most dispersed and internally differentiated of the African language phyla (Map 1). The first monograph on the phylum is Schadeberg (1981), followed by Bender (1996) and Ehret (2001). Recent research on the ‘green Sahara’ has suggested that 12,000 years ago the desert was threaded through with interconnected waterways (Drake p.c.), explaining the fossil remains and representations of aquatic creatures such as hippos occurring across the region. This also corresponds neatly to the distribution of bone harpoons and in part to the distribution of Nilo-Saharan languages (Map 1).

---

4 It is very striking how this formulation does not change over the decades, despite the publication of numerous reference books clearly adopting ‘Iger-Congo’ as a reference name.
It is therefore suggested that the origins of Nilo-Saharan lay in its area of greatest diversity, SW Ethiopia, but that its stimulus to expansion was the opening up of a major new resource, the waterways of the Sahara.

### 4.1.4 Afroasiatic

In the mid-1990s, two very different perspectives on the phylum were published, both accompanied by substantial bodies of data, Ehret (1995) and Orel & Stolbova (1995), comparative lexicons of Afroasiatic with proposals for sound correspondences. Strikingly, these voluminous studies propose very different internal structures for Afroasiatic and a list of protoforms at odds with one another. Figure 2 shows a proposed internal classification of Afroasiatic.

### Figure 2. Internal structure of Afroasiatic

In Blench (2006) it is proposed that the possibility of reconstructing ‘cow’ in all but Omotic, argues for a primary split between Omotic vegeculture and Cushitic intensified animal management, eventually developing into true pastoralism after several millennia. This would suggest that reconstructions of
agriculture are spurious as are suggestions that Afroasiatic expanded out of the Near East and is associated with the Natufian (e.g. Militarev 2002).

4.2 Eurasia

4.2.1 Dravidian

Our understanding of Dravidian is strongly related to the Dravidian etymological dictionary of Burrow and Emeneau (1984 and online). However, this is very Tamil-centric and the literature constantly confuses its head entries with proto-Dravidian (e.g. Krishnamurti 2003). Notwithstanding these reservations, Southworth (2005, 2006) has undertaken an analysis of this data in terms of subsistence reconstructions with generally convincing results. Broadly speaking, the earliest phase of Dravidian expansion shows no sign of agriculture but (lexically) reflects animal herding and wild food processing. This is associated with the split of Brahui from the remainder. The next phase, including Kurux and Malto, shows clear signs of agriculture (taro production but not cereals) and herding, while South and Central Dravidian have the full range of agricultural production. Fuller (2002) and Southworth (2006) link this to the aptly named South Neolithic Agricultural Complex (SNAC) dated to around 2300-1800 BC in Central India.

4.2.2 Indo-European

Indo-European is one of the most researched of the language phyla of the world, and perhaps this has a history of deadening debate on its origins. Figure 3 shows a typical ‘tree’ with all the major subgroups placed on an equal footing.

Figure 3. Classification of the Indo-European languages

The essential debate consists of those who associate PIE with horse-culture in the steppes of Central Asia (Anthony 2007) and those who link it to the expansion of agriculture from Anatolia (Renfrew 1987). For a linguist there should be no contest; horses and other livestock species have deep-level reconstructions in Indo-European, whereas words unambiguously denoting crops are confined to subsets of branches. The proposal here is that the early Indo-European speakers were managers of wild horses, gradually bringing them and small ruminants under control. They fished and exploited crab-apples and only later adopted agriculture from their near-Eastern neighbours.

4.2.3 Uralic

The Uralic phylum consists of some twenty-four languages spoken from Hungary to Siberia. The phylum may include the Yukaghir language is spoken on the Kolyma river in NE Siberia. Collinder (1960, 1965), Redéi (1968-9) and Décsy (1990) have all compiled etymological dictionaries of the phylum. Hájdu (1975, 1976) lists the etymologies relevant for the subsistence strategies of proto-Uralic speakers. He dates the original expansion of the phylum to 8-6th millennia BP in the North-Central Urals, basing his hypothesis on a combination of palaeoecological data and reconstructions of names for tree species. Terms for fishing and hunting can be reconstructed, as well as terms suggesting a northern habitat. More problematic are his reconstructions of domestic animals, such as ‘sheep’ and ‘pig’. These are based principally on Hungarian and Finnish evidence, and their absence from Samoyedic argues that they are not to be reconstructed to the

5 See http://www.suri.ee/eup for an excellent summary of the current condition of individual Uralic-speaking groups
proto-language. Uralic is thus another example of a foraging expansion, like Pama-Nyungan. Its expansion need not be attributed to any major motivating force, but the slow expansion of populations into very difficult terrain, through the evolution of hunting techniques.

4.2.4 Altaic [?]

The minimal set of Altaic languages consists of Turkic, Mongolic and Tungusic, spread from Turkey to Siberia. Most scholars accept that Korean also forms part of Altaic and some that Japanese is also a member. There have been proposals to also include Ainu, but this is not generally accepted. Starostin et al. (2002) have published a major etymological dictionary of comparative Altaic, which provides rich material for interpreting its prehistory. Compared with the other phyla discussed here, Altaic has a very unusual substructure. Its branches are internally very close-knit, and the dispersal of Turkic has largely taken place in historical time, but very different from one another, so much so that some scholars claim that Altaic is not a phylum but a bundling of languages that have interacted (Janhunen 1994).

Mongolic languages today are dominated by Khalkh Mongol, spoken throughout much of modern Mongolia, with outlying Mongolic languages spoken in China and Afghanistan (Janhunen 2003). The relative uniformity of Mongolic can be attributed to the empire of Chinggis Khan (ca. 1200 to 1400 AD) which grew to control the largest land empire ever recorded and probably eliminated earlier ethnic and linguistic diversity during this period. Janhunen (1993) has analysed lexical elements borrowed from Mongolic into Manchurian Tungusic to argue that the family formerly exhibited much greater diversity.

Apart from Manchu, the Tungusic languages all have a small number of speakers whose populations were until recently hunter-gatherers. Starostin (2002) points out that this may be a reversion to hunting and gathering as Tungusic shares names for crops with other Altaic languages. Surprisingly, however, the Tungusic languages are not highly diverse compared with other Siberian populations, suggesting that the Tungusic expansion is recent. It seems likely that Tungusic groups were spread more widely across northern Heilongjiang Province and the adjacent Mongol-speaking area, probably in interaction with Koreanic speakers. It is considered possible that Tungusic speakers were responsible for the introduction of the Northern Bronze Complex into the Korean peninsula during the 1st millennium BC, and also that the Rong people, associated with the Upper Xiajiadian in south-eastern Mongolia, represent a southern intrusion of Tungusic (Barnes 1993:165).

4.2.5 Kartvelian

Kartvelian was previously referred to as South Caucasian, as if were to be paired with North Caucasian, but this name has been dropped as it was realised that these language phyla have in fact no special relationship. The Kartvelian language family comprises the Georgian, Megrelian, Laz and Svan languages spread throughout a large territory to the south of the Transcaucasian mountain range. Kartvelian must be a relatively recent expansion, since its members are all relatively close to each other. Klimov’s (1998) etymological dictionary of Kartvelian provides a rich source for reconstructions of the proto-language. Table 1 lists the proto-Kartvelian terms relating to agriculture;

<table>
<thead>
<tr>
<th>Proto-Kartvelian</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>*anc̃l̃-</td>
<td>elder tree</td>
</tr>
<tr>
<td>*bag-</td>
<td>paddock</td>
</tr>
<tr>
<td>*beɣ-</td>
<td>barn</td>
</tr>
<tr>
<td>*burč̃x-</td>
<td>*Echinochloa crus galli</td>
</tr>
<tr>
<td>*cel-</td>
<td>to scythe, scythe</td>
</tr>
<tr>
<td>*čal-</td>
<td>he-goat</td>
</tr>
<tr>
<td>*ćičil̃-</td>
<td>chicken</td>
</tr>
<tr>
<td>*ćša</td>
<td>rice grains in husk</td>
</tr>
<tr>
<td>*ćuk-</td>
<td>cauldron</td>
</tr>
<tr>
<td>*ćwel-</td>
<td>straw</td>
</tr>
<tr>
<td>*ćwel-</td>
<td>to milk</td>
</tr>
<tr>
<td>*dab-</td>
<td>field, village</td>
</tr>
<tr>
<td>*daq- (*dqa-)</td>
<td>goat</td>
</tr>
</tbody>
</table>
Proto-Kartvelian  | English
--- | ---
*dil̞-* | wheat
*dol-* | winter wheat
*ešw | boar, pig
*gen-* | calf; to suck (udder)
*yom-* | gomi (a k. of millet)
*yor-* | pig
*ɣryad-* | goose
*ɣ̝wino-* | wine
*ipkL-* | wheat
*kat-* | hen
*kum-* | oatmeal
*kwrim-* | a k. of millet
*krav-* | lamb
*ku̞pr-* | wine-skin
*kwic-* | mountain goat
*la̞yw-* | fig

This demonstrates that the proto-Kartvelians were cultivators, and indeed agriculture must go back in this region much further than the 2-3000 years required for the internal diversity of the phylum.

### 4.2.6 North Caucasian

North Caucasian consists of some thirty-four languages spoken by some 4.6 million people in modern Russia, Georgia and Azerbaijan. It is also claimed that the extinct languages, Hurrian-Urartian and Hattic-Khaldi belonged to this phylum. The phylum is sharply divided into the Northeast and Northwest groups and not all researchers believe these constitute a linguistic unity. Nikolayev & Starostin (1994) have published an etymological dictionary of North Caucasian which they believe began to split up > 5000 years ago. They claim that terms for agriculture can be reconstructed to this hypothetical proto-language, but it is difficult to speculate on such an uncertain genetic entity.

### 4.2.7 Sino-Tibetan [=Tibeto-Burman]

Sino-Tibetan is the phylum with the second largest number of speakers after Indo-European, largely because of the size of the Chinese population. The internal classification of Sino-Tibetan remains highly controversial, as is any external affiliation. The key questions are whether the primary branching is Sinitic (i.e. all Chinese languages) versus the remainder (usually called Tibeto-Burman), the position taken by Matisoff (2003) and most other scholars (Thurgood & LaPolla 2003) or is Sinitic simply integral to existing branches such as Bodic etc. as Van Dreim (1997) has argued. It is certainly true that the established ‘Sino-Tibetan’ model has barely been supported by linguistic arguments (though see Bodman 1980).

It is therefore not reasonable at present to reconstruct the history of Tibeto-Burman through either internal genetic classification or comparative lexicron. At present, we can only go by internal diversity and there is no doubt that this is greatest in the I epal-Bhutan area. The present assumption is that the diverse groups were originally hunter-gatherers making seasonal forays onto the Tibetan Plateau but that 7-6000 BP this became permanent occupation, probably due to the domestication of the yak (Aldenderfer & Yinong 2004; Xuebien et al. forthcoming).

### 4.2.8 Andamanese

Andamanese languages are confined to the Andaman Islands, west of Myanmar in the Andaman Sea. The Andamanese are physically like negritos, i.e. they resemble the Orang Asli of the Malay peninsula and the Philippines negritos and ultimately Papuans. It is unlikely that Andamanese forms a linguistic phylum. No data exists on Sentinelese, and Great Andamanese appears not to be related to Onge and Jarawa, which perhaps are related (Blench 2008).
4.2.9 Austroasiatic

Our understanding of Austroasiatic has been much increased by the publication of Shorto’s (2006) comparative dictionary. Proto-Austroasiatic speakers almost certainly already had fully established agriculture. It is possible to reconstruct ox, ?pig, taro, a small millet, numerous terms connected with rice and ‘to hoe’, all with Muṇḍā cognates. If so, it seems that Austroasiatic may be ‘younger’ than the time-scales proposed by Diffloth (1997, 2005). However, our understanding of Austroasiatic is limited by the lack of material on Pakanic and other more remote branches which may represent its earliest phases and such sites may therefore represent a later expansion.

4.2.10 Daic [=Tai-Kadai, Kradai]

The Daic or Tai-Kadai languages, of which Thai is the most well-known and widespread representative, are spoken from southern Thailand into Laos, Cambodia and China. Overviews of the phylum are given in Edmondson & Solnit (1988, 1997a). Figure 4 shows the internal classification of Daic according to Edmondson & Solnit (1997b);

![Figure 4. Internal classification of Daic](image)

Daic languages are not all that diverse and almost certainly a candidate for a major agricultural expansion. Despite this, there is no obvious archaeological correlate. Blench (2005) has presented some evidence for thinking that speakers of proto-Daic were not originally rice-cultivators, that they borrowed these techniques from Austroasiatic speakers. Reconstruction has yet to produce evidence for their subsistence strategies, and it may be that they were originally cultivators of tubers such as taro, which would fit with the links with the islands. But without a deeper knowledge of the pattern of Daic dispersal it is hard to link them directly with
any of the known archaeological horizons of south China. Ostapirat (2005) argues that Daic and Austronesian are genetically related, a demonstration accepted by many scholars and Sagart (2004) has placed Daic on a branch corresponding to Malayo-Polynesian. If so, then the proto-Daic speakers would have migrated back from the southern tip of Taiwan about 4000 BP.

4.2.11 Miao-Yao

The Hmong-Mien [=Miao-Yao] languages are spoken mostly in China with some groups also in Laos, Vietnam and Thailand. Their centre of gravity is between the Yangzi and the Mekong rivers. Hmong-Mien languages are quite close to one another, and although the Ethnologue lists some 32 languages, many of these are mutually intelligible lects. There have been various comparative overviews of the group, starting with Purnell (1970), Wang Fushi (1994), Wang & Mao (1995) and Iiederer (1998). Tapp et al. (2004) have edited a comprehensive overview of recent scholarship including much valuable bibliography. The linguistic geography of Miao-Yao suggests very strongly that these people were scattered by the incoming Han and probably forced southwards into modern Laos and Thailand, probably in the last 3-4000 years. This has sparked a number of debates on the relative antiquity of these groups; if Miao-Yao preceded Chinese, should it not be more diverse? Sagart (p.c.) has put forward the suggestion that pre-Miao was adopted by the Chu state (500 BC onwards) which would have had a Sinicised bureaucracy. The northern distribution of Miao probably represents the boundary of this state. Yao, the more southerly languages, must have escaped this state at some stage and were perhaps within another state, as Yao languages have a unique set of Sinitic loans. Ratliff (2004) has explored the reconstructible environmental and agricultural vocabulary in Miao-Yao.

4.3 Pacific

4.3.1 Austronesian

Austronesian is the second-largest language phylum in the world after Niger-Congo and certainly the most widespread, stretching from Easter Island to Madagascar. Figure 5 shows the high-level structure of Austronesian according to Blust (1999);

Figure 5. High-level structure of Austronesian according to Blust (1999)

![Diagram of Austronesian language structure]

Austronesian may be considered almost a holotype for an agricultural expansion (Blust 1995; Bellwood, Fox & Tryon 1995; Pawley 2002). Its speakers probably left the Chinese mainland 6-7000 BP for Taiwan already with some type of cultivation. They spent more than 1000 years on Taiwan and then very suddenly, just before 4000 BP, began expanding with great rapidity south to the northern tip of the Philippines. Pawley (2002:257) says, ‘the third striking feature is the swiftness of the spread of *Neolithic* traditions after they reached the Philippines’. Indeed the metaphor of an ‘express train’ is regularly used in the literature (e.g.
Diamond 1988). Among other achievements, they reached the Marianas islands in the middle of the Pacific at this period which indicates a very advanced maritime technology. This is reflected in the controversial classification of Malayo-Polynesian, which unlike the Oceanic languages appears to split into a number branches whose relation to one another remains ill-defined.

4.3.2 Papuan

Papuan is not usually thought of as a single phylum; the present view is that to consists of the Trans I ew Guinea (Tl G) and a number of small families and isolates (Pawley 1995, 2005). The exact number of these varies between authors; Foley (1986) thought there were some sixty distinct families. Ross (2005) lists twenty-three including the Tl G. However, the Trans I ew Guinea phylum would be by far the largest grouping in I ew Guinea and its expansion is therefore the most relevant to the present discussion. The ‘Indo-Pacific’ hypothesis of Greenberg (1971) which unites Tasmanian, Papuan and Andamanese has regrettably found its way into many archaeologically –oriented texts has never been validated by linguists. A recent review by Pawley (n.d.) suggests that Greenberg had some insights into an early version of the Tl G hypothesis, but that the links with Andamanese and Tasmanian were entirely spurious.

4.3.3 Trans New Guinea

Pawley (2005) reviews the history of the Trans I ew Guinea phylum, which began life as the ‘East I ew Guinea Highlands’ stock proposed by Stefan Wurm. Over the years its membership has expanded and it is now considered to cover most of the island of Papua I ew Guinea (Wurm & McElhanon 1975; maps in Wurm & Hattori 1981, Ross 2005:34). Debates continue about the affiliation of particular language subgroups, but there appears to be general acceptance of the core hypothesis. Lexical cognacy rates between remote branches can be barely above levels of chance resemblance, but nonetheless, it has proven possible to list around 200 putative reconstructions. Pawley (2005:97) considers that internal lexical diversity points to a date of at least 10,000 BP and perhaps earlier. It has been proposed that agriculture was the primary motivation for the expansion of the Tl G and this would indeed fit with the dates for taro cultivation, for example at Kuk swamp (Denham 2007). However, it faces the objection that apart from the widespread #ma root for ‘taro’ (which also occurs outside the Tl G), there are virtually no items suggesting agriculture in the reconstructed lexicon of Tl G. Pawley (2005:98) points to the importance of arboriculture and also the changes in geomorphology which would have changed the shape of the island over the last ten thousand years. It may well be that, despite the sporadic presence of agriculture, it was intensive tree management that powered the expansion of the Tl G. However, documentation of individual tree names is not sufficiently well advanced to make possible the reconstruction of particular species.

4.3.4 Australian

‘Australian’ represents a convenient cover term for the indigenous languages of Australia, but should not be taken to imply their genetic unity. It is widely estimated there were some 400+ languages in Australia prior to European contact, and that of these records remain for at least 280 (Dixon 2002). Ethnologue (2005) estimates that some 260 languages are still spoken although many may now only have one or two old speakers. Australian languages have considerable similarities in terms of phonological and grammatical subsystems; however, it has not proven possible to unite them in single genetic group, although some linguists still consider this a possibility (Koch 2004). The ‘Common Australian’ referred to in works by Dixon (e.g. Dixon 1980) usually reflects the characteristics of Pama-I yungan. Figure 6 represents a simplified tree of Australian languages, remaining agnostic about the possibility that all Australian languages are related;
4.3.5 Pama-Nyungan

Most of Australia is covered by a single phylum, Pama-Nyungan, with all the remaining languages confined to the north. It also implies, given the time-depth of human settlement in Australia, that the expansion effectively eliminated a far greater linguistic diversity if the previous situation was anything like the existing non-Pama-Nyungan phyla. Pama-Nyungan has been the subject of a reconstruction project which may eventually recover several thousand common roots (Fitzgerald 1997; O’Grady 1998; Koch 2004). Pama-Nyungan is also the holotype of a non-agricultural phylic expansion. While the role of agriculture can be debated elsewhere in the world, its complete absence in Australia shows that it can have played no role in the expansion of Pama-Nyungan. Pama-Nyungan clearly expanded relatively recent and reflects either a technological advantage or a strikingly different social system (McConvell & Evans 1998). The usual date assigned to Pama-Nyungan is just 4,5000 BP.

4.4 New World

4.4.1 Eskimo-Aleut

Eskimo-Aleut covers a vast region of the subarctic zone from Siberia to eastern Canada. The grouping of the two languages was first mooted in 1818 by Rasmus Rask. Eskimo is usually assigned a time-depth of about 2000 years and the difference with Aleut suggests the original split was at least two thousand years before that (Dumond 1987). Dumond connects the expansion of Eskimo-Aleut with the ‘orton tradition’ of check-stamped pottery. All Eskimo-Aleut remain foragers and there is no clear evidence what technical innovation they possessed which permitted their spread across such a swathe of inhospitable terrain. Fortescue et al. (1994) have published a comparative Eskimo-Aleut dictionary.

4.4.2 Na-Dene

The dominant Indian language family of northwestern America is Na-Dene, a name given by Edward Sapir (1915) deriving from the Haida and Athapaskan words for ‘people’. This family consists of Haida, Tlingit, Eyak, and Athapaskan. The Athapaskan family is a complex of some 30 languages and dialects. Athapaskan has spread far beyond its original area of some 2000 years ago, perhaps in the upper Yukon River region, throughout much of interior Alaska and northwestern Canada, and thence to southern Oregon and northern California, and separately in the Southwest, where it is spoken by the Navajo and Apache. Na-Dene is problematic (Dürr & Renner 1995). Scholars such as Campbell (1997) do not accept that Haida is part of Na-Dene and use the reduced form. Enrico (2004) presents evidence for the affiliation of Haida, but also accepts that there are many early loanwords that make the evidence problematic. Vajda (2008) has presented evidence for a link with the Yeniseian languages of Siberia which seems to have gained wide acceptance, which case the phylum may be renamed Dene-Yeniseian. The entire phylum consisted of foragers until
recently, so exactly what might have impelled its dispersal is difficult to suggest. If the deep connections with Siberia are accepted then it is likely that Dene-Yeniseian is an old phylum (perhaps up to 10,000 years BP) although the Athapaskan extension may be as late as 2000 years ago.

4.4.3 Mayan
The Mayan languages form a phylum with 30+ members spoken in Mexico and Guatemala. Speakers of the lowland Mayan languages have brought fame to the family as a whole through their spectacular stone architecture and their writing system. Campbell (1997:165) mentions the agricultural inventory of Mayan specifically and extensive cognate sets can be found in Dienhart (1989). Brown and Wichmann (2004) have proposed a long list of reconstructions for crops and a variety of terms connected with agriculture. A date of > 4000 BP is usually attributed to proto-Maya. However, whether the adoption of agriculture was the motivating factor for Mayan expansion remains open to doubt, since this is also a group with highly elaborate political and social systems which can also be reconstructed back to the era of the proto-language.

4.4.4 Otomanguean
The Otomanguean languages are spoken in a wide swathe of territory between southern Mexico and (formerly) Nicaragua. Longacre & Millon (1961) reconstructed proto-Amuzgo-Mixtecan specifically with a view to identifying subsistence modes of its speakers. Rensch (1976, 1989) reconstructed a large number of crop names for proto-Otomanguean and Table 2 shows reconstructions relating to proto-Otomanguean subsistence that have been proposed by various authors.

Table 2. Proto-Otomanguean subsistence reconstructions

<table>
<thead>
<tr>
<th>Authors</th>
<th>Crops</th>
<th>Other terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longacre &amp; Millon (1961), Rensch (1976)</td>
<td>avocado, bean sp., cacao, chili, maguey, maize, sweet potato (or <em>camote</em>?), squash, cotton, tobacco</td>
<td>maize dough, <em>metate</em>, <em>pulque</em>, oven, <em>pulque</em></td>
</tr>
</tbody>
</table>

Hopkins (1984) has connected the spread of Otomanguean with the rise of agriculture in the Tehuacán Valley (The Tehuacán tradition is a horizon 5000-2300 BC).

4.4.5 Mixe-Zoque
The Mixe-Zoque (MZ) languages are spoken principally in Southern Mexico around the Isthmus of Tehuantepec. Campbell & Kaufman (1976) argued that proto-MZ could be identified with Olmec civilisation and indeed that Olmec script could be deciphered with a knowledge of the reconstructed proto-language. This direct equation has been questioned by Olmec scholars such as Michael Coe, but there is clearly some relationship between the spread of Olmec culture and the Mixe-Zoque. Wichmann (1995) describes the internal classification of Mixe-Zoque and Wichmann (1998) considers agricultural vocabulary. Although there is little doubt that agricultural terms can be reconstructed for proto-MZ, there is clearly also considerable borrowing and internal loaning.

4.4.6 Uto-Aztecan
Uto-Aztecan is a family of languages stretching between the southern United States and Southern Mexico and including the language of the Aztecs. Earlier arguments supposed that this was originally a forager phylum e.g. Fowler (1972, 1983), Miller (1984) but Bellwood (2001, 2005 and elsewhere) and others (e.g. Hill 2002) have turned this argument on its head and claimed it as an agricultural expansion from Meso-America into the southwestern US. This however, would involve ‘devolution’ i.e. the return of at least some populations (I umic-speakers) to foraging. Campbell (2002) argues strongly that the linguistic evidence is very weak and archaeological evidence so far non-existent.

4.4.7 Cariban
The Cariban language phylum is widespread across northern South America, from the mouth of the Amazon River to the Colombian Andes and from Maracaibo (Venezuela) to Central Brazil (Figure 7). Cariban languages are relatively close to each other with 20-30 still spoken. Villalón (1991) has made a strong case for the Cariban expansion as being essentially ‘trading and raiding’ rather than agricultural although an absence of published reconstructions means that the interpretation of Carib prehistory is rather inconclusive. Nonetheless, the presence of outlying languages along major rivers rather than in the forest tends to support
this approach. Meira & Franchetto (2005) show that the Southern groups are quite coherent and there is no argument for a southern origin of Cariban based on diversity.

Figure 7. The Cariban languages

4.4.8 Arawakan [=Maipuran]

The Arawakan languages are spoken from the eastern slopes of the central Andes Mountains in Peru and Bolivia, southward into Paraguay and northward to the north coast of South America (Aikhenvald 1999). Their extension into the Caribbean is thought to be recent. Arawakan is the largest family in the Americas with the respect to number of languages and covers the widest geographical area. It has been the subject of a number of puzzlingly contradictory linguistic reconstructions (e.g. Noble 1965; Matteson 1972; Payne 1991) partly because the affiliation of languages such as Arauan and Harakmbet is unsettled. Payne’s version includes at least some agricultural terms, but the evidence remains tenuous. The cultural reconstruction of Arawakan has a long history, with a first attempt by Schmidt (1917). Williams (2003) has proposed a link between the so-called ‘Timehri’ petroglyphs and Arawak expansion and Heckenberger (2002) a cultural model that links them to the ‘Barrancoid’ ceramics. These models may have a partial validity but the correlations do not cover the whole Arawakan-speaking area. Hornborg (2005) has proposed ditching the ‘traditional’ migrationist model in favour of ‘modern’ ethnogenesis. While this may play well in anthropology circles, it seems very unlikely to be true in the Amazon, a vast space cut through with waterways, providing almost ideal conditions for actual migration.

5. Synthesis

5.1 The pattern of isolates and small families

Looking at the worldwide pattern of isolates, it is evident that they are very unevenly distributed (Table 3). There is almost a gradient from west to east, with few in Europe and the greatest number in the I ew World. To compare like with like, known Eurasian isolates that have long become extinct, such as Sumerian and Etruscan are excluded. The high density of isolates in the Americas, is surely no accident but tells us something very important about the peopling of the I ew World, namely that such a rich diversity cannot have arisen within the constrictions of the chronology accepted by many I orth American archaeologists. For so many languages to have been diversifying for so long as to eliminate all traces of links with neighbouring languages requires time-depths similar to those accepted for Papua and Australia.
Table 3. Isolates, small phyla by continent

<table>
<thead>
<tr>
<th>Continent</th>
<th>No. Isolates</th>
<th>No. Small Phyla</th>
<th>Total living languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>6</td>
<td>0</td>
<td>2092</td>
</tr>
<tr>
<td>Eurasia</td>
<td>6</td>
<td>1</td>
<td>2508</td>
</tr>
<tr>
<td>Pacific</td>
<td>12</td>
<td>4</td>
<td>1079</td>
</tr>
<tr>
<td>Australia</td>
<td>7</td>
<td>13</td>
<td>263</td>
</tr>
<tr>
<td>New World</td>
<td>70</td>
<td>43</td>
<td>1002</td>
</tr>
</tbody>
</table>

In some ways the enterprise of the long-rangers is very misleading, since their efforts to sweep all the worlds’ languages into a set of macrophyla has the effect of actually obscuring the different between major linguistic regions. The pattern of the Americas is no accident of linguistic history but a very real difference from Africa or Australia for reasons that are yet to be determined.

Table 4 shows the accepted language phyla of the world and contrast them with ‘groupings’, geographical ensembles such as Papuan. Very small phyla and isolates are excluded. [?] implies controversy in the literature about the status of a phylum.

Table 4. Language phyla and groupings of the world

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Phylum Grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khoesan</td>
<td>+</td>
</tr>
<tr>
<td>Niger-Congo</td>
<td>+</td>
</tr>
<tr>
<td>ilo-Saharan</td>
<td>+</td>
</tr>
<tr>
<td>Dravidian</td>
<td></td>
</tr>
<tr>
<td>Andamanese</td>
<td>+</td>
</tr>
<tr>
<td>Papuan</td>
<td>+</td>
</tr>
<tr>
<td>Trans New Guinea</td>
<td>+</td>
</tr>
<tr>
<td>Australian</td>
<td>+</td>
</tr>
<tr>
<td>Pama-Nyungan</td>
<td>+</td>
</tr>
<tr>
<td>Tasmanian</td>
<td>+</td>
</tr>
<tr>
<td>Afroasiatic</td>
<td>+</td>
</tr>
<tr>
<td>Indo-European</td>
<td>+</td>
</tr>
<tr>
<td>Uralic</td>
<td>+</td>
</tr>
<tr>
<td>Altaic [?]</td>
<td>+</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>+</td>
</tr>
<tr>
<td>Orth Caucasian</td>
<td>+</td>
</tr>
<tr>
<td>Sino-Tibetan</td>
<td>+</td>
</tr>
<tr>
<td>Austroasiatic</td>
<td>+</td>
</tr>
<tr>
<td>Austronesian</td>
<td>+</td>
</tr>
<tr>
<td>Daic</td>
<td>+</td>
</tr>
<tr>
<td>Miao-Yao</td>
<td>+</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>+</td>
</tr>
<tr>
<td>a-Dene</td>
<td>+</td>
</tr>
<tr>
<td>Mayan</td>
<td>+</td>
</tr>
<tr>
<td>Otomanguean</td>
<td>+</td>
</tr>
<tr>
<td>Arawakan</td>
<td>+</td>
</tr>
<tr>
<td>Tupian</td>
<td>+</td>
</tr>
<tr>
<td>Mize-Zoque</td>
<td>+</td>
</tr>
<tr>
<td>Uto-Aztecan</td>
<td>+</td>
</tr>
<tr>
<td>Cariban</td>
<td>+</td>
</tr>
<tr>
<td>Cariban</td>
<td>+</td>
</tr>
<tr>
<td>Panoan</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 5 shows accepted language phyla and the evidence (or lack of it) for subsistence. This table excludes all ‘geographical’ groupings. Citing a reference does not imply I accept the authors’ conclusions;
### Table 5. Language phyla of the world and linguistic evidence for subsistence

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Evidence exists</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niger-Congo</td>
<td>+</td>
<td>Westermann (1927), Blench (2006)</td>
</tr>
<tr>
<td>Ilo-Saharan</td>
<td>+</td>
<td>Bender (1996), Ehret (2001), Blench (ined.)</td>
</tr>
<tr>
<td>Dravidian</td>
<td>+</td>
<td>Southworth (2005)</td>
</tr>
<tr>
<td>Trans-New Guinea</td>
<td>—</td>
<td>Pawley (2005)</td>
</tr>
<tr>
<td>Indo-European</td>
<td>+</td>
<td>Numerous</td>
</tr>
<tr>
<td>Altaic (?)</td>
<td>—</td>
<td>Starostin et al. (2002)</td>
</tr>
<tr>
<td>Kartvelian</td>
<td>+</td>
<td>Klimov (1998)</td>
</tr>
<tr>
<td>I orth Caucasian (?)</td>
<td>—</td>
<td>I kologayev &amp; Starostin (1994)</td>
</tr>
<tr>
<td>Austroasiatic</td>
<td>+</td>
<td>Shorto (2006)</td>
</tr>
<tr>
<td>Austronesian</td>
<td>+</td>
<td>Blust (ined.)</td>
</tr>
<tr>
<td>Daic</td>
<td>—</td>
<td>Ostapirat (2000)</td>
</tr>
<tr>
<td>Miao-Yao</td>
<td>+</td>
<td>Ratliff (ined.)</td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>+</td>
<td>Fortescue et al. (1994)</td>
</tr>
<tr>
<td>I a-Dene</td>
<td>—</td>
<td>Enrico (2004)</td>
</tr>
<tr>
<td>Arawakan</td>
<td>+</td>
<td>Payne (1991)</td>
</tr>
<tr>
<td>Uto-Aztecan</td>
<td>+</td>
<td>Hill (2002)</td>
</tr>
<tr>
<td>Carib</td>
<td>+</td>
<td>Durbin (1977), Villalón (1991)</td>
</tr>
<tr>
<td>Tupian</td>
<td>+</td>
<td>Lemle (1971)</td>
</tr>
</tbody>
</table>

Needless to say, it is only my prejudices that dismiss published evidence for subsistence in some cases.

### 5.2 Phyla: dates, homelands, motives for expansion

Table 6 then shows the accepted language phyla of the world with estimated dates for dispersals, possible homelands and possible ‘motors’ of expansion. For many phyla, controversies between different researchers have meant that no unambiguous engine of expansion can be identified. African phyla are based on my own proposals.
<table>
<thead>
<tr>
<th>Phylum</th>
<th>Date BP</th>
<th>Homeland</th>
<th>Engine</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>† Ilo-Saharan</td>
<td>&gt;15,000</td>
<td>SW Ethiopia, Uganda</td>
<td>Climate improvements, fishing</td>
<td>‘Green Sahara’</td>
</tr>
<tr>
<td>† Niger-Congo</td>
<td>&gt;10,000</td>
<td>Southern margins of the Sahara</td>
<td>I ew hunting techniques</td>
<td>Ounanian?</td>
</tr>
<tr>
<td>Afroasiatic</td>
<td>&gt;10,000</td>
<td>SW Ethiopia</td>
<td>Livestock management</td>
<td></td>
</tr>
<tr>
<td>Dravidian</td>
<td>&gt;5000</td>
<td>Central India</td>
<td>Livestock management</td>
<td></td>
</tr>
<tr>
<td>Trans I ew Guinea</td>
<td>&gt;10,000</td>
<td>I ew Guinea Highlands</td>
<td>Intensive arboriculture, Agriculture ?</td>
<td>Kuk swamp etc.</td>
</tr>
<tr>
<td>Pama-I yungan</td>
<td>&gt;4000</td>
<td>I ort hern Australia</td>
<td>Lithic technology, religion?</td>
<td></td>
</tr>
<tr>
<td>Indo-European</td>
<td>&gt;6000</td>
<td>Central Asia</td>
<td>Horse pastoralism</td>
<td></td>
</tr>
<tr>
<td>Altaic</td>
<td>&gt;8000</td>
<td>I E Asia</td>
<td>Horse pastoralism</td>
<td></td>
</tr>
<tr>
<td>Kartvelian</td>
<td>&gt;3000</td>
<td>Georgia</td>
<td>Agriculture, political organisation</td>
<td></td>
</tr>
<tr>
<td>I orth Caucasian</td>
<td>&gt;5000</td>
<td>Caucasus</td>
<td>Agriculture [?]</td>
<td></td>
</tr>
<tr>
<td>Sino-Tibetian</td>
<td>&gt;7000</td>
<td>Himalayas</td>
<td>Pastoralism [?]</td>
<td></td>
</tr>
<tr>
<td>Austroasiatic</td>
<td>&gt;10,000</td>
<td>South China [?]</td>
<td>Taro cultivation [?]</td>
<td></td>
</tr>
<tr>
<td>Austroenesian</td>
<td>&gt;6000</td>
<td>Taiwan</td>
<td>? Maritime technology</td>
<td></td>
</tr>
<tr>
<td>Daic</td>
<td>&gt;4000</td>
<td>China</td>
<td>Agriculture</td>
<td></td>
</tr>
<tr>
<td>Miao-Yao</td>
<td>&gt;3000</td>
<td>China</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Eskimo-Aleut</td>
<td>&gt;3000</td>
<td>I E Asia</td>
<td>Foraging</td>
<td></td>
</tr>
<tr>
<td>I a-Dene</td>
<td>&gt;5000</td>
<td>I E Asia</td>
<td>Foraging</td>
<td></td>
</tr>
<tr>
<td>Maya</td>
<td>&gt;2500</td>
<td>Yucatán Peninsula</td>
<td>Agriculture, political organisation</td>
<td>Tehuacan valley horizon</td>
</tr>
<tr>
<td>Otomangueean</td>
<td>&gt;3000</td>
<td>Tehuacan valley, Mexico</td>
<td>Agriculture, political organisation</td>
<td></td>
</tr>
<tr>
<td>Arawakan</td>
<td>&gt;4000</td>
<td>Orinoco Basin?</td>
<td>Foraging, trading?</td>
<td>Barrancoid ceramics, Timehri petroglyphs ?</td>
</tr>
<tr>
<td>Cariban</td>
<td>&gt;4000</td>
<td>Orinoco Basin?</td>
<td>Foraging, trading?</td>
<td></td>
</tr>
<tr>
<td>Mixe-Zoque</td>
<td>&gt;4000</td>
<td>Isthmus of Tehuantepec</td>
<td>Agriculture, political organisation</td>
<td></td>
</tr>
<tr>
<td>Uto-Aztecanc</td>
<td>&gt;4000</td>
<td>Southern Mexico</td>
<td>Foraging</td>
<td></td>
</tr>
<tr>
<td>Cariban</td>
<td>&gt;4000</td>
<td>Orinoco Basin?</td>
<td>Foraging</td>
<td></td>
</tr>
</tbody>
</table>

6. Conclusions
The literature on phylic expansions has been dominated by solutions proposed by archaeologists, and agriculture is to hand as a simple motivating factor for demographic increase and thus language diversification. For one type of publication there is an ‘Emerging Synthesis’ (e.g. Renfrew 2002). But linguistic support for this is very slight, and mostly for small and more recent phyla only. Forager spreads, such as † Ilo-Saharan, † Niger-Congo, Pama-I yungan and Eskimo-Aleut, cover the greatest geographical area, and the engines of their expansion remain debated. The internal structure of many language phyla,
including Indo-European remain uncertain. Even for the holotypes, Austronesian and Bantu, long accepted by sceptics as principally motivated by agriculture, doubts can be raised. The example of Pama-Nyungan shows that very large-scale and significant expansion/assimilation processes can occur solely within the framework of foraging societies, and that we accept the likelihood of this occurring in regions of the world that predominately agricultural today.

At root, the argument about the significance of agriculture is a reframing of Childe’s Neolithic revolution combined with Sauer’s emphasis on demography. It sees the transition to agriculture as the most important transformation that can occur within the evolution of a language phylum. But as evidence gathers for the slow speed and gradualness of such transitions and the effectiveness of intensive landscape management it becomes clearer that agriculture is only one factor among many in the evolution and dominance of individual language phyla. Moreover, it is also the case that agriculture is likely to play a significant role somewhere in the subsets of languages that make up a phylum; in other words, if the argument does not appear to work for the phylum as a whole, apply only to a subset. This substantially weakens the explanatory power of the model.

One possible solution is to invert the argument, to suggest that agriculture can be a consequence rather than a cause of demographic change. In other words, if a population is expanding, perhaps due to resource abundance, medical innovation, new military or political behaviours or technology, then there is a greater incentive to make the transition from intensive plant management to agriculture proper. In African models, this would be the change from cultivation to domestication, from animal management to pastoralism. In these terms, the key Austronesian innovation would not have been agriculture, even though the speakers were already cultivators, but maritime technology. Once the Austronesians invented extremely effective outriggers suitable for long-distance ocean voyaging, the proto-Malayo-Polynesians began to disperse all through the region. As a population scattered across the islands of SE Asia and the Pacific, agriculture began to dominate their resource extraction repertoire. In the case of Niger-Congo, the initial spread may have been both through the hunting of large game with the newly-adopted bow and arrow and paddling down the large rivers exploiting aquatic environments. When population began to exceed foraged resources, first cultivation and then domestication became the proximate solution to a resource crisis. Early Dravidian, for which only livestock terms can be reconstituted, may have reflected mobile pastoralism or animal management. Only later, when other expanding language phyla place pressure on foraged resources does agriculture become dominant.

This argument can be summarised as follows;

a) Language goes back into the unknown past of foraging societies and possibly to the inception of modern humans. The default behaviour of such societies is to slowly expand demographically, and their languages to eventually diversify to a point where individual speech-forms are no longer relatable to one another.

b) A major ‘punctuation’ occurs when there is a change in resource availability, or the technical or social capacity to exploit that resource. These factors underlying these changes can be external, such as climate change, or internal, such as religious innovation or the invention of the outrigger or the bow and arrow.

c) Such changes provide a significant impetus to the ethnolinguistic group affected by them and it expands geographically, either demographically or through assimilation. This process is more effective among foragers than among cultivators but may result in fewer languages. Agriculture can have the effect of slowing down language distatiñation, although there may be great numbers of due to the increased in speakers.

d) The consequence is a pattern of geographically extensive language phyla dispersed among isolates or small phyla. Such extensive phyla can be fragmented or coherent, depending on the nature of the impulsion.

e) Gradual intensification of plant or animal management, to the point where it can be defined as agriculture, may therefore occur when the expanding phylum encounters a resource bottleneck. Where there is no bottleneck, foraging continues, and where the resources/demography equation favours foraging, devolution back to foraging can occur.
f) This explains why terms definitely indicative of agriculture can often only be reconstructed to a part of individual language phyla (and why a coherent model of the internal structure of individual phyla is essential to the reconstitution of language prehistory)

Figure 8 attempts to present these processes in the form of a diagram;

**Figure 8. Foraging, the transition to farming and language diversification**

---

It is difficult not to be struck by the importance of waterways and aquatic subsistence in early language phylum expansions, both maritime and inland. This may be no accident, as recent hypotheses concerning primary human expansion out of Africa suggest coastal migrations (e.g. Stringer 2000). Illo-Saharan, parts of Iger-Congo, Arawakan, Cariban and Austronesian may well all have been aquatic expansions with agriculture evolving to support the consequent demic expansion. This is not to try and introduce another model to replace the significance of agriculture; many language phyla clearly have little or nothing to do with waterways, such as Pama-I yungan. However, it is worth trying to rebalance the model to consider that a wide range of factors may have played their part in what remain dramatic processes.

There is another significant consequence for linguists; the reconstruction of language phyla is too important to be left to them. This may seem perverse; it is usually argued that disciplines should proceed in isolation and reconstruct proto-languages based on purely linguistic considerations. Only then can their results be compared with data from other sources such as archaeology, genetics or palaeoclimatology. However, this has not really proven very successful in practice; even the most well-known phyla have disputed structures and doubtful reconstructions. In part this is because if a linguist really knows nothing about historical ecology, or archaeological results, they are likely to make errors of assumption and fact, such as reconstructing New World plants for Old World phyla. If we are better informed about the climate and ecology of the relevant region as well as the chronological stratification of agriculture or other distinctive subsistence strategies, we are better placed to make credible reconstructions, and to better distinguish early loanwords from true cognates. Historical linguistics may thus actually depend on history rather than existing in a vacuum of abstraction.
References


**Websites**

http://www.alaskool.org/language/languageindex.htm