

Why Is Africa So Linguistically Undiverse? Exploring Substrates and Isolates

Roger Blench

*Kay Williamson Educational Foundation, Cambridge, U.K.*¹

For the special issue of *Mother Tongue* celebrating the work of Joseph Greenberg

1. Introduction

One of the notable differences between Africa and most other linguistic areas is its relative uniformity. With few exceptions, all of Africa's languages have been gathered into four major phyla, and most recent progress in classification has been in resolving details. The number of undisputed language isolates is very small. By contrast, Australia, Papua and the New World are usually considered diverse at the phylic level and all have numerous isolates or very small phyla. Eurasia is hard to classify; Europe is undiverse and is characterised by a small number of geographically extensive languages, but NE Asia is a reservoir of small phyla. SE Asia, on the other hand, is very similar to Africa, in having few phyla, each with many languages and no isolates, if Andamanese is excluded.

Looking at the worldwide pattern of isolates, it is evident that they are very unevenly distributed. There is a gradient from west to east, with few in Europe and the greatest number in the New World. Probable Eurasian isolates that are long extinct, such as Sumerian and Etruscan, point to a period of greater diversity, but crucially, they have been assimilated. The high density of isolates in the Americas tells us something very significant about the proposed chronology of the peopling of the New World, namely that such richness cannot have arisen within the chronological constraints accepted by many North American archaeologists (Blench 2012). 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 1 shows the different regions of the world and the numbers of isolates and small phyla, by my own count, but based on sources such as *Ethnologue* (Lewis et al. 2013), Campbell (1997a) and Pawley (2012). Totals do not include extinct languages, otherwise the numbers for the New World would be significantly higher. Total numbers of living languages come from *Ethnologue*:

¹ Kay Williamson Educational Foundation, 8, Guest Road, Cambridge CB1 2AL, United Kingdom | Voice / Ans (00-44)-(0)1223-560687 | Mobile worldwide (00-44)-(0) 7847-495590 | E-mail rogerblench@yahoo.co.uk | <http://www.rogerblench.info/RBOP.htm>

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

Table 1. Isolates, small phyla by continent

Continent	No. Isolates	No. Small Phyla	Total living languages
Africa	6	0	2146
New World	71	43	1060
Eurasia	6	1	2588
Pacific	12	4	1066
Australia	7	13	245
Total			7105

This view of diversity conflicts radically with the ‘long-ranger’ view of the world language phyla, whereby the New World can be rolled up into just three families and much of Melanesia into a single family, Indo-Pacific. The major issue is whether this reflects intellectual tradition or linguistic reality which is considered below. However, it should be noted here that Greenberg (1971, 1987) simply does not discuss the classification of many languages in the New World and Melanesia now claimed to be isolates by regional specialists.

Given the time-depth of human settlement in Africa, this situation of uniformity is somewhat surprising. If the *ex Africa* hypothesis for the origin of modern humans is accepted, then *Homo sapiens sapiens* originated some 150-200 Kya and spread to Eurasia from Northeast Africa, largely displacing, but perhaps also interbreeding with, the hominids already *in situ*. It is now widely accepted that modern humans evolved in Africa (Ke *et al.* 2001). The first evidence for archaic modern humans is in the Omo Valley some 195 kya ago (Macdougall *et al.* 2005). This are likely to have been three waves of hominin dispersals out of Africa between 1.9 and 0.7 Ma (Bar-Yosef and Belfer-Cohen 2001). Genetic and archaeological studies suggest that anatomically modern humans (AMH) dispersed out of Africa sometime between 130 and 50 ka (Gunz *et al.* 2009). Behar *et al.* (2008) have demonstrated that there was a major division in human populations some 70 kya leading to the genetic isolation of the Khoisan in Southern Africa. This must be connected with the finding that the ancestor of modern males, associated with mitochondrial haplotypes M and N, appears at this time. The claim is that M168 mutation was carried out of Africa and is characteristic of all non-African males. They associate this with the megadrought in Eastern Africa between 135 and 70 kya (Cohen *et al.* 2007).

The routes by which modern humans left Africa remain disputed. Cave sites in Israel at Qafzeh make it certain that one exit was via Sinai into the Near East. In the early 2000s, it was commonly accepted that modern humans left Africa via the Bab el Mandeb, the modern Horn of Africa, crossed to Arabia and followed the coastlines all the way down to Australia. This scenario found support from DNA analyses of Andaman Islanders and Orang Asli (Macaulay *et al.* 2005). Another likely dispersal route is across the Sahara, either via the Nile (Vermeersch 2001) or the ‘green Sahara’ during more humid periods in the past (Drake *et al.* 2011). There is evidence for AMH occupation of North Africa including the North African littoral, the Sahara and the Nile in the form of the Nubian and Aterian lithic industries (Van Peer 1998). These archaic humans, or similar peoples, could

have dispersed out of Africa at around 120 ka (Dennell and Petraglia 2012) in the light of archaeological remains in the Levant (Grün et al. 2005) and the existence of Nubian technology in Arabia (Rose et al. 2011). However, a coastal is yet to be confirmed by excavation, despite considerable advances in the archaeology of Arabia (Petraglia 2007; Petraglia & Rose 2009). Current views are tending towards a model whereby foragers crossed the centre of the Arabian Peninsula at a period when the environment was considerably more favourable (Petraglia et al. 2010).

Given the antiquity of human evolution in Africa, the low number of isolates and lack of linguistic diversity clearly requires an explanation. A number of hypotheses can be advanced to account for this. The simplest is that languages diversify at non-uniform rates and therefore there is no puzzle. This is hard to disprove at some level, since it can be demonstrated for some historical cases. For example, the evolution of script systems appears to have a major impact on slowing rates of language diversification, as does the growth of state systems. However, none of these seem to be relevant in the present case, where these areas have remained populated by small, ethnolinguistically distinct groups without writing. If we accept there is a link between apparent diversity and time-depth, essentially three alternatives remain:

- a) Africa was once as linguistically diverse as other regions of the world and this diversity has been eliminated by the expansion of the ancestors of today's phyla since the Holocene
- b) Intellectual traditions concerning classificatory processes are significantly different in other regions of the world and the apparent disparity is illusory, e.g. Papuan, Australian and Amerind do form coherent, so far unrecognised phyla
- c) *or* the classification of African languages is radically in error

Joseph Greenberg, whose classification of African languages undoubtedly formed the basis for all subsequent classifications, certainly focused on the second explanation. His arguments for Indo-Pacific (Greenberg 1971) and Amerind (Greenberg 1987) claim that most of the languages of these regions *can* be classified into large phyla². The third alternative, that we are simply wrong about classification, can be associated with Africanist linguists such as Gerrit Dimmendaal (2011), Tom Güldemann (2008, 2011) and the global linguist Harald Hammarström³, who suggest that Africa has many more small families than are currently accepted, and that we have failed to recognise this. Importantly, this is a claim that Greenberg (1963) is seriously in error. However, it is problematic to discuss their

² Although these ideas had precursors, notably the observations of Trombetti (1905: 205-208) who noted the prevalence of 1st person N- and 2nd person M- pronouns in numerous American languages, Greenberg was the first to take the bolder step of claiming these were a key piece of evidence for a continent-spanning phylum. Greenberg (1996) cites some of his predecessors in this area. The notion that these pronoun distributions are statistically significant has been questioned by various authors, notably Campbell (1997b) and Brown (n.d.) who compares Amerind pronominal patterns with worldwide distributions.

³ Hammarström's views are summarised at <http://glottolog.org/glottolog/family> which certainly represents the extreme end of the spectrum. Not only is Nilo-Saharan discounted, but major branches within it are dismantled. Omotic is similarly dismembered and several branches of Niger-Congo treated as isolated families.

argument, since by and large Dimmendaal and Güldemann are not interested in historical linguistics. They assert that Mande or Songhay have been misclassified but do not present any argument against the traditional view.

This paper⁴ assumes that Greenberg's overall picture for Africa was broadly right, but that his arguments for other linguistic regions of the world were at best over-optimistic. This is particularly the case for the Americas, where no recent papers by regional specialists have supported Amerind, and many have rejected it. The only supporting arguments appear in the work of Merritt Ruhlen⁵ (e.g. 1994, 1995). Indo-Pacific is more complex, since some of the languages Greenberg grouped together *are* now considered to be part of the Trans New Guinea phylum (Pawley 2005, 2012). Whitehouse et al. (2004) further include the Nepalese isolate Kusunda within Indo-Pacific⁶. Andamanese and Tasmanian have failed to garner further support in the intervening period, despite a major expansion of data on Andamanese (Abbi 2006, 2012). Crowley & Dixon (1981) considered the evidence for a Tasmanian link within Indo-Pacific but came up with strongly negative conclusions. The overall assessment of the research community in relation to Greenberg's proposals can be measured by the establishment of scholarly conference series, which exist for Niger-Congo, Nilo-Saharan, Afroasiatic and Khoisan (i.e. all the African groups) but not for either Amerind or Indo-Pacific.

This paper examines the alternative explanations and seeks a synthesis to account for the current situation. However, it also asks questions about the modelling of linguistic diversity more generally and by inference how we assign chronologies and time-depths to phyla with no significant written records. This is clearly not just an artefact of the way we do linguistics, but somehow reflects a real phenomenon. The final section of this paper will consider why it should be that Africa, generally considered to be the original home of modern humans, should have such low linguistic diversity.

2. Methodological issues

2.1 Traditions of classificatory research in Africa and beyond

The perceived diversity of a linguistic region is not entirely the result of a rigorous scientific process; it also reflects strongly the patterns established in the early period of scholarship. In African studies, the intellectual tradition has been characterised from an early period by continent-spanning hypotheses. The discovery that Bantu languages from

⁴ One section of this paper was first presented in Canberra on 30th July 1999 in the Department of Archaeology and Anthropology. A more complete version was circulated at the workshop 'Language Isolates in Africa' Lyon, December 3 and 4, 2010. For more recent materials on African isolates I'd like to acknowledge the help of Philippe Beaujard, Harald Hammarström, Abbie Hantgan, Maarten Mous, and Martin Walsh in making available and commenting on various materials. Work on data analysis was supported by the Kay Williamson Educational Foundation.

⁵ Ruhlen re-edited the Amerind etymological dictionary after Greenberg's death (Greenberg & Ruhlen 2007).

⁶ Kusunda was thought to be extinct, but surprisingly some speakers were contacted in 2004 and a grammar and wordlist have now been published (Watters 2005). These do not provide support for the link with TNG. Blench (2008) notes that Kusunda has a distinctive vocabulary for agriculture, suggesting they are not archaic hunter-gatherers, but a reversion from a farming culture.

Cameroun to South Africa were related dates back to the seventeenth century⁷ (Doke 1961) and the discovery of noun-classes in West African languages led some nineteenth century scholars to speculate on their relation to Bantu. Wilhelm Bleek (1862, 1869) included a West African division in the family he named Bantu. Meinhof, one of the most influential figures in early twentieth century African linguistics, was originally a Bantuist and it is more than possible that the 'large-pattern' world view that this engendered influenced much of his later work, now considered misconceived, that linked together numerous languages in his 'Hamitic' hypothesis. Diedrich Westermann (1911) posited a 'Sudanic' family corresponding to Meinhof's work on Bantu, and proposed a division between 'East' and 'West' Sudanic, linking together what we would now consider Nilo-Saharan and Niger-Congo. Westermann's second major publication (1927) set out a large number of Proto-West Sudanic (PWS) reconstructions, broadly corresponding to modern Niger-Congo and playing a role analogous to the work of Dempwolff (1934-1938) in Austronesian reconstruction.

Joseph Greenberg took a fresh look at the classification of African languages in a series of articles published between 1949 and 1954, later collected in book form in 1963. He combined 'West Sudanic' and Bantu into a phylum he named Niger-Congo, while he treated 'East Sudanic' as a different phylum, renamed Nilo-Saharan. He renamed the 'Hamito-Semitic' languages Afroasiatic and while re-iterating the Khoisan hypotheses of Dorothea Bleek (1956), who had assumed not only that all the click languages were related to one another, incorporated the languages with clicks in East Africa, a hypothesis later named Macro-Khoisan. The effect of this was to tidy up the linguistic picture of the whole continent –every language was theoretically 'placed'. Why Greenberg seems not to have considered the possibility of isolates is unclear, but his approach has been enormously influential on succeeding generations of Africanist scholars. Indeed, Greenberg's later publications, first on Indo-Pacific (Greenberg 1971), then gathering all the languages of the Americas into three phyla (Greenberg 1987) and bringing together Eurasian languages into 'Eurasian' (Greenberg 2000), a version of Nostratic, show that he was a committed 'lumper'.

These views remain very much at odds with more conventional scholarly opinion on the languages of Papua, Australia and the New World. In these regions, linguists have generally entered the field with no preliminary assumptions about relatedness or macro-groupings –and so progress has been much more 'bottom-up'. Small groups have been derived from data and gradually built into larger ones. Pama-Nyungan and the Trans-New Guinea Phylum (TNG) represent 'entry-level' attempts to try and discern larger patterns, but they are far from encompassing the whole repertoire of languages and are still treated with scepticism in many quarters⁸. Similarly, the Amerindianist tradition shares much in common with Papuan. The earliest classifications, such as that of John Wesley Powell in

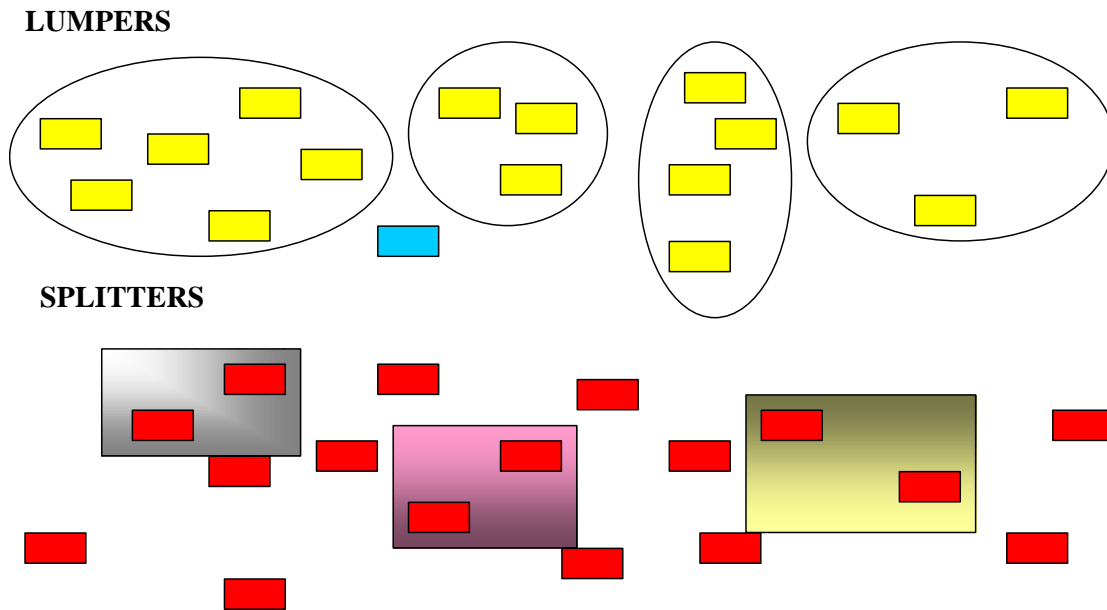
⁷ Though an intriguing passage in Strabo, Geography Book II, Chapter 3 about the Greek navigator Eudoxus of Cyzicus (fl. c. 130 BC) suggests that the similarities of Bantu may have been noticed much earlier.

⁸ See Pawley (2005, 2012) for a re-evaluation of the TNG.

the 1890s, divide Amerindian languages into numerous unrelated families. Although Amerindianists have gradually been reducing this number⁹, they have essentially worked from the assumption that languages constitute distinct phyla until they are shown to be related (Campbell 1997; Campbell & Poser 2008). But, unlike Greenberg's proposals for Africa, Amerind and other constructs have gained very limited assent from the scholarly community. *Language in the Americas* was reviewed by a variety of scholars in *Current Anthropology* 28.5 (1987) with rather mixed results. More tellingly, the individual datasets were reviewed by numerous specialist scholars in the early 1990s in the *International Journal of American Linguistics*. Like the headline review (Rankin 1992), almost all were resoundingly negative. Greenberg (1996) mounted a spirited defence of Amerind claiming in effect that a small number of inaccuracies did not constitute a disproof of the hypothesis. However, a point made by many discussions is the inadmissible segmentation of words, leading to pseudo-cognates, i.e. claiming incorporated morphology in one language is cognate with a stem in another, and there is no obvious response to this. More recent discussions have not necessarily improved the picture (e.g. Rankin 2012).

The early tradition within a region is thus extremely powerful in determining the pattern of phyla, families and isolates. If Africa were in Melanesia, as it were, its linguistic geography might well be a few larger phyla and many isolates characterised by complex contact phenomena. Figure 1 is a schematic model of lumpers and splitters which tries to visualise the impact they have on our perception of the linguistic geography of a region.

Figure 1. Schematic model of lumpers and splitters



⁹ Not always. Arawakan has 'lost' languages since earlier reconstructions, since it was realised Harakmbet is unrelated.

Is the conclusion that we might also be wrong about Africa? That Nilo-Saharan, Niger-Congo or Khoisan are no more than networks of isolates, or much smaller phyla, and the supposed cognates simply borrowings or chance. If we depended only on Greenberg's 'mass comparison' this might indeed be the case, since it is now all too apparent how significant borrowing can be between languages. But with the exception of Khoisan, our existing phyla are probably safe, because Greenberg depended on a body of prior work. Although he rarely cites it, it provided much of the evidence underlying his proposals. The comparative nominal morphology of Niger-Congo was first laid out in detail by Westermann (1935) and it is largely one of the accidents of history that this is not regularly cited as the key paper in establishing the phylum. Recent publications (beginning with Dixon 1997 but characterising some of the papers in Heine & Nurse 2008) and conferences, such as 'Beyond Niger-Congo' have used geographical and typological mapping of traits to suggest that Niger-Congo in particular is somehow not a valid phylum. This type of argument is problematic for two main reasons. First the authors rarely engage with the literature, failing to describe the errors that presumably characterise the proposals of those who want to argue for the reality of particular phyla. But more important it represents a methodological error, the assumption that demonstrating contact phenomena or mapping typological traits constitutes an argument against genealogical affiliation.

As an example of this type of construct, consider 'Macro Sudanic' (Güldemann 2008, 2011). This consists of a series of maps and tables showing that particular phenomena (labial-velars, logophoricity, vowel harmony) have quite similar distributions across a wide area of northern Sub-Saharan Africa. Güldemann concludes from this that the historical linguists are wrong and that 'the Macro-Sudan belt is genealogically highly heterogeneous'. In other words, this is a reprise of the arguments of David Dalby, who plotted a 'Fragmentation Belt' across Africa. But evidence that widespread phenomena are partly attributable to genealogical characteristics of language and partly to contact does not discredit historical linguistics. Indeed it is hard to imagine what type of scholar would not accept such a model. The point, and it is an important one, is that some linguistic phenomena are more prone to diffusion than others.

Even one of Güldemann's own examples, the labial flap, does not seem a very convincing demonstration of his point. The labial flap has a rather discontinuous distribution within Africa, found in some languages and not others. As Olson & Hajek (2003) show, there are 'islands' in Southern African Bantu, very remote from the main region of distribution. How are these to be interpreted? Was the labial flap part of proto-Bantu and has simply died out elsewhere? This seems somewhat unlikely. Did it arise independently? Given its global rarity, again improbable. It turns out that the labial flap is far more common than is suggested in the 2003 survey (see additional material in Anonby 2007). Matthew Harley (p.c.) reports it in several languages in central and northwest Nigeria, including Tərin [Pongu]. As an example of how it can be simply not heard because it is unexpected, take the example of Bafut, a Grassfields language. Despite its clear presence, previous accounts of Bafut phonology fail to recognise it (Mfonyam 1989). The likely explanation of its presence in Shona is that it *is* scattered across Bantu; not a

reconstructible phoneme, but carried by marginal lexicon, for example in ideophones, and simply not recognised in many linguistic descriptions. This also exemplifies the issue of defects in the descriptive literature. Güldemann maps the distribution of \pm ATR vowel harmony as it was known around 2000. However, it turns out that much of the literature on Nilo-Saharan was inaccurate, and that, for example, the Kadu languages and Fur also have vowel harmony. Linguistic geography is highly contingent; it depends on the phenomena you decide to map, the literature you consult, and the state of linguistic description, which itself may reflect politics and financial resources. It has little to do with the argument about whether cognate morphemes in Niger-Congo affixing systems constitute proof or otherwise of the reality of the phylum¹⁰.

2.2 Misclassification of marginal languages

Any reasonably complex linguistic situation tends to present a mosaic of languages, some that show strong links with their neighbours and those that are less easy to classify. At least in Africa, such difficult to classify languages have a correlation with small, remote, marginal or isolated groups. The implicit assumption of most historical linguists has been to group the clearly-related languages, name them and then assume that the others must fit in somewhere. The evidence for this has all too often been a rather small number of common lexical items –an approach pioneered by Greenberg in ‘mass comparison’ and rehearsed in all too many documents since.

To give some concrete examples, the Fali languages of Northern Cameroun were classified by Greenberg as a group of Adamawa Eastern (now Adamawa). Lexical, grammatical and survey data is available for all these languages (Sweetman 1981a,b) and Fali has no single lexical or morphological feature that links it unambiguously to Adamawa, although it ‘ought’ to be Adamawa in terms of linguistic geography. However, the inertia of classification is such that no author has come out and simply asserted that it is ‘not’ Adamawa. Another language misclassified by Greenberg as Adamawa, Chamba Daka, has been excised and is generally considered Bantoid (Boyd 1994). This type of negative argument is much harder to make than one for a positive affiliation and probably less intellectually exciting, especially if you have no alternative proposal. Similar are the Ega language of Cote d’Ivoire, regularly cited as ‘Kwa’ despite a maximum 9% lexicostatistic resemblance to other purportedly related languages and Seme, in Burkina, Faso, which shows a similarly low level of resemblance to Kru.

A parallel can be drawn here with the classifications of Indo-European. We tend to think of Indo-European as a widespread phylum with a small number of well-known languages, Latin, Greek, English, Sanskrit, etc. However, there are well over 400 Indo-European languages, the majority of which are spoken in India. Some of these have intriguing and unusual features such as elaborate tone systems and complex phonology, as well as exhibiting considerable lexical diversity. However, they are all classified as ‘Indo-

¹⁰ Larry Hyman (2011) has also presented a detailed critique of Güldemann’s methods and results, although using very different examples from those given here.

Aryan' on the grounds that 'tribal' languages are merely 'dialects' of the larger languages and that those in turn can be related back to Sanskrit (e.g. Masica 1993). There is no material at all on the sources of the non-standard vocabulary. These languages could just as easily be the non-Indo-European speech of resident hunter-gatherers relexified through contact with the vernaculars of incoming agriculturalists. However, the pattern of scholarship and the non-availability of key data on specialised lexicon makes this speculation difficult to pursue.

3. Language isolates in Africa and elsewhere

The list of African isolates remains controversial, depending on whether those with apparent substrates are counted. There is also one interesting case of a language which appears to be spurious (Oropom). Table 2 lists the languages that have remained doubtfully classified and Map 1 shows their locations:

Table 2. African language isolates

Language Name	Location	Source	Comments
Jalaa (=Cun Tuum)	Nigeria	Kleinwillinghöfer (2001)	Probably extinct
Bangi Me	Mali	Blench (2007a), Hantgan (p.c.)	
Laal	Chad	Boyeldieu (1977), Lionnet (2010)	
Kujarge	Sudan	Doornbos & Bender (1983); Blench (2013); Lovestrang (2012)	Probably East Chadic
Ongota	Ethiopia	Fleming (2006), Savà & Tosco (2000)	Perhaps Afroasiatic
Hadza	Tanzania	Sands (1998), Miller (p.c.)	
Sandawe	Tanzania	Sands (1998)	Probably Khoesan
Kwadi	Angola	Westphal (1963), Güldemann (2004)	Perhaps Khoesan. Probably extinct

Of these, it seems likely that Kujarge and Sandawe are respectively Afroasiatic and Khoisan, giving a total of six.

Map 1. Africa: languages isolates and residual foragers



There are number of languages which have been reported initially as isolates, but which seem to be affiliated to known phyla, as shown in Table 3:

Table 3. African language reported isolates

Name	Location	Source	Comments
Bēosi	Madagascar	Blench (2007b), Blench and Walsh (n.d.)	Austronesian with unknown ? Southern Cushitic substrate
Dompō	Ghana	Painter (1967), Blench (n.d. b)	Guang language with unknown substrate
Gumuz	Ethiopia	Bender (1979), Ahland (2010)	Nilo-Saharan isolate branch
Mbre	Cote d'Ivoire	Creissels (n.d.), Blench (n.d. c)	Niger-Congo isolate branch
Mpra	Ghana	Cardinall (1931), Blench (n.d. a)	Extinct. Kwa language
Oropom	Uganda	Wilson (1970)	Probably spurious
Shabo	Ethiopia	Bender (1983), Fleming (1991)	Nilo-Saharan isolate branch

4. Substrates and isolates: their definition and detection

4.1 Where does substrate vocabulary show up?

If it is the case that Africa was once substantially more diverse, do traces of this diversity persist? In particular, if foragers became farmers and lost their original language in the process, under what conditions should some elements of it remain in their speech? Global evidence for the persistence of substrate vocabulary is very mixed. Evidence from place names (including those recorded in Ptolemy's *Geography*, 2nd Century AD) suggests that British Isles was once populated by populations speaking a language of unknown affiliation, Pictish (Jackson 1955). Some inscriptions in the Ogham script are untranslatable (Lockwood 1975) while others are clearly Celtic. But very few items of substrate vocabulary remain in English, and these can be identified principally by our comparative knowledge of surviving Celtic speech and literary records. In regions of the world where we have no such records, fragmentary vocabulary such as this would be extremely difficult to identify. By contrast, classical Greek is considered to have just 22% of its lexicon traceable to Indo-European roots, and the remainder presumed to derive from pre-Greek languages, 'Pelasgian', which were almost certainly non-Indo-European (Heubeck 1961; Beekes 2010).

It is therefore likely that these differences are related to the sociolinguistics of language assimilation. Powerful states which impose their languages on new subject populations are more likely to insist on language replacement. Slow language shift is more likely to produce lects which retain the lexicon and sometimes the phonological and morphological structures of previous languages. A good example of this in Africa is Yei, a Bantu language spoken in the Caprivi Strip. It retains or has borrowed a substantial number of 'click' words, especially to do with plants and animals, characteristic of the Khoesan languages in this region which it must have replaced (Sommer & Voßen 1992).

Interestingly, these words retain their original click status rather than being phonologised; hence Yei has a very large inventory of rare sounds.

The detection of substrates requires, above all, a rich descriptive literature. Substrates are likely to survive in specialised lexical fields, for example animal and plant names, hunting and fishing technology, religious terms. It almost always will also be present in morphology and syntax, but these are always much harder to interpret. For example, in the Kadu languages of Kordofan, a relatively coherent, closely related set of languages, two of the nine lects are SOV instead of SVO (Hall & Hall 2004; Blench 2006a). This is surprising, given that most African language families are highly conservative about word order. Is this change in Kadu due to the influence of a substrate or bilingualism in neighbouring languages? Both word orders occur in the region. Or is just an internal change, part of language evolution? The arguments are lengthy and often unresolved, which is generally not the case with distinctive technical vocabulary.

4.2 Chance? Can we exclude fortuitous resemblances

The identification of isolates depends on the tools used to classify languages. If a language shows only a small number of problematic cognates with its proposed relative, then its genetic affiliation will inevitably be questioned. Nilo-Saharan and Khoisan in particular include languages whose inclusion in the phylum remains debated. Several of the languages of the Ethio-Sudan borderland, such as Shabo and Gumuz within Nilo-Saharan (e.g. Bender 1979) and the 'Mao' languages, particularly Ganza, within Omotic, have very low lexical cognate count with their relatives (Bender 2003). If we claim that a substrate in a language can be identified in the lexicon of a quite different language, what counts as proof of lexical resemblance? Three explanations are possible:

- a) the putative branches have been diverging away from the rest of the phylum for sufficiently long for vocabulary erosion to be responsible for low lexical counts
- b) apparent similarities with the other branches of the phylum are due to borrowing
- c) chance

Linguistic analysis, the demonstration of regular sound-correspondences or the detection of loanword phonology should be sufficient to show whether a) or b) are probable. But what about chance? There is a literature suggesting that lexical lists of any two languages in the world might show up to 5% resemblances of CVC stems (Bender 1969). Calculations by Ringe (1992, 1999) have applied a great deal of energy to algorithms illustrating the difficulties of showing languages are related. So the suggestion that the resemblances leading to a proposal of a relationship are 'chance' appears at first sight persuasive.

But in fact attributing resemblance to 'chance' is a virtually worthless heuristic, because it is an untestable proposition, since no empirical data can ever be adequate to exclude it. Amassing evidence may make any linguistic proposition more likely, but a negative can never be demonstrated. In other words, it can never be shown that the apparent relation between two lexemes is *not* due to chance. Clearly, it is always possible to find

unrelated languages where individual items show close sound/meaning correspondences. Our assumption that the languages in question are unrelated is partly determined by geography, partly by the lack of a regular relationship. But the regularity of a relationship can really only be determined by comparative data. If one language shows lookalikes and its genetic relatives do not, borrowing or chance may be the explanation. But if languages have no close relatives, then it is problematic to exclude these alternatives. The calculations by Bender and Ringe assume that languages have no structure, that in principle any combination of CV phonemes may arise. But in practice this is not true, since most languages are extremely constrained in their permissible canonic structures. If two languages are related, then the set of lexemes said to be cognate should have constraints on both phonology and canonic forms.

The assumption of chance is thus an unusable tool. We can draw up tables of more or less likely cognates, and whether these are accepted by other linguists is a function of the credibility of the sound-meaning correspondences and demonstration that these are not borrowings. The reduction of the numbers of phyla in Australian languages is a good example of this. Earlier descriptions supposed there were at least twenty-five Australian language phyla, a figure which has been reduced to ten or eleven as a consequence of a better understanding of historical morphology. This in turn is a function of the growth of descriptive literature and the intensity of research on these languages. African languages are never likely to have the same degree of attention; results come slowly. However, they do come, as witness the characterisation of Omotic (Bender 1975, 2001, 2003) and its gradual acceptance by the Afroasiatic establishment.

4.3 Nilo-Saharan and historical linguistics

The previous section mentions widespread doubts over the affiliation of languages claimed to be Nilo-Saharan. Part of the issue is the way historical linguistics is conventionally conducted. Compared with other African phyla, the difficulties of demonstrating the reality of Nilo-Saharan have typically propelled authors into methodological excursions (Bender 1997; Ehret 2001; Blench 2002). It seems highly unlikely that Nilo-Saharan will ever pass the tests of regular sound correspondences and an agreed internal structure now part of the formula for the usual textbooks on historical linguistics. In other words, Nilo-Saharan will never look like Austronesian or Dravidian. There are simply not enough undisputed lexical cognates to set up secure correspondences or develop clouds of isoglosses illustrating particular subgrouping hypotheses. For those who really don't want to see Nilo-Saharan there will probably can never be enough evidence to support it.

Greenberg (1987), prescient as ever, was one of the first to point to the problems with textbook historical linguistics. His preferred method, 'mass comparison', was intended to circumvent the issues of poor data and low-level cognacy. However, whether 'mass comparison' could ever be treated as a method rather than as the insights of an exceptional individual is open to doubt. Greenberg's 'success' with African languages is often cited as a reason for accepting his more controversial hypotheses, such as Eurasiatic

or Indo-Pacific. But the truth is that Nilo-Saharan is the only one of his African hypotheses that was genuinely new. Niger-Congo, Afroasiatic and Khoisan were all picked up from previous authors (Westermann (1927), Cohen (1947), Bleek (1956)) and neither Amerind nor Eurasiatic have received comparable scholarly assent¹¹. Nonetheless, in the grand scheme of things, Greenberg was the first to outline numerous points of lexicon and morphology that remain of importance.

The problem arises because the methods we use for historical linguistics derive from 'neater' phyla. A phylum that has gradually expanded across a terrain (Austronesian, Uralic) and falls into well-defined subgroups is always going to be closer to this ideal structure. But the historical conditions responsible for Nilo-Saharan are quite different. It apparently dispersed across Africa in a period when foraging was still the only method of subsistence. A recent model proposes that this was an east-west expansion and that the key factor was the availability of aquatic resources in Saharan and Sahelian regions (Drake, Blench et al. 2011). Later expansions of unrelated phyla have encapsulated or heavily influenced each of its branches, inducing heavy lexical and structural borrowing and extensive morphophonological shifts. History has often moved on, obscuring these patterns. The consequence is that lexical cognates are few, fragmentary and may now be widely geographically dispersed, to the extent that they can be hard to distinguish from chance.

However, the historical morphology of Nilo-Saharan plays a role in making cognates so difficult to discern. Niger-Congo languages typically have morphological features which characterise most of its branches, most notably nominal class affixes and verb extensions. Although these do not occur everywhere, they are widespread and when lost can often be recognised through fossil morphology. Indeed, as Westermann (1935) pointed out long ago, the detailed similarities between the affixes across widely geographically disparate branches provide useful preliminary evidence for the reality of the phylum.

Nilo-Saharan has no such instantly recognisable features. Some branches have nominal class affix alternation (e.g. Koman, Daju, Kadu) but these do not show alliterative concord and the affixes are not obviously linked with semantics. Languages of the Kadu group are particularly misleading, since at first they seem to have nominal affixes, leading Greenberg (1963) to mistakenly consider 'Tumtum' as part of Kordofanian (Blench 2006). But these are very remote from the systems of Niger-Congo, and form part of a complex three-term system of noun plurals. I have previously argued that Nilo-Saharan must once have had a noun-class system like Niger-Congo and hence the two phyla could be united

¹¹ John Bengtson (p.c.) comments 'The only scientifically valid way to refute Greenberg's Indo-Pacific or Amerind is to show that a different classification is superior'. However, demonstrating negative hypotheses, i.e. that two languages are not related, is not the structural opposite of a positive claim. The specialist may claim that the proposed phonological relationships or semantic shifts are 'unlikely' but we have no gold standard of what is unlikely. Clearly improbable shifts do occur. The linguistic community will inevitably go with the weight of scholarly opinion, in the awareness it may be revised.

into a singular macrophylum (Blench 1995). I now consider that the common features of the two phyla can equally well be explained by early borrowing.

Nilo-Saharan has characteristics which make recognising cognates particularly challenging. These are:

- a) words erode from the front, i.e. they lose C_1 [usually a morpheme] and sometimes replace it with another consonant [also often a morpheme]¹²
- b) metathesis of the root is common (which is probably related to a)
- c) conservation of unproductive affixes (as many as three on one word) making it difficult to discern the root.

The underlying reason is probably an original numeral classifier system resembling Asian languages, which is heavily eroded or appears synchronically only as frozen affixes. Various studies have noted associations between affixes and semantic themes. For example, both Stevenson (1991) and Gilley (2014) note the semantic associations of affix pairings in Kadu languages, and Storch (2005) remarks on these for Western Nilotic. But the most striking evidence comes from Gumuz, whose Mayu dialect has been studied by Ahland (2010). Gumuz has a system of nominal incorporation, where a series of body part nouns are incorporated into verbs and which 'classify' the object, or more rarely the subject or instrument. Blench (2014) argues that this reflects the original system of Nilo-Saharan, where the predicate classifiers are bound to the verb, or a system of nominal classifiers preceding the noun, which were appended to the verb in Gumuz. Ahland (2010) reports that the neighbouring Bertha language may also have such a system, although this has not been written up. If this is the case, it explains the disconcerting way number markers are copied between prefix and suffix slots in Nilo-Saharan. Gumuz shows clearly that a range of permissible word orders allow the predicate classifier to show up in different places in the sentence.

The typological parallel here is Austroasiatic, a phylum with thirteen branches like Nilo-Saharan, and also dispersed through aquatic movement, in this case along the Mekong some 4000 years ago (Sidwell & Blench 2011). Austroasiatic has a system of roots and single consonant prefixes, although the prefixes have no obvious semantic associations synchronically. However, they can vary alarmingly from one branch of Austroasiatic to another, and to establish cognates it is necessary to ignore them and compare stems. Fortunately, C_1 can be deleted or replaced in Austroasiatic, but the prefix does not become a suffix, nor does the prefix combine with the stem consonant to produce complex consonants, all of which make it possible to recognise cognates more easily.

The methodological point should now be more evident. Some languages at first sight look problematic to relate to others, because of our rather rigid views of historical morphology. There is a sort of psychodrama in linguistics between the 'long-rangers',

¹² One of the curious aspects of this morpheme replacement is that they appear to have no synchronic meaning in any Nilo-Saharan branch. Greenberg and Bender both assumed they were articles, which makes sense, but cannot be supported from descriptive data.

whose house-journal is *Mother Tongue* and the distinct but related the website 'Tower of Babel' established by Sergei Starostin. Those who see long distance connections between languages are often impatient with the sceptics and some can be seen as riding roughshod over the usual rules of historical linguistics. At the other end of the spectrum are those who can never relate anything to anything else unless rigorous sound-correspondences are in place, represented by Glottolog. Nilo-Saharan is not made for them. We have to tread a line between Voltairean genealogies, where everything is related, and out and out scepticism, accepting that some phyla and macrophyla mean bending the rules.

5. Detecting substrates in animal names

5.1 Madagascar

One of the most interesting scenarios for retention of substrate vocabulary is where an expanding population comes into contact with a new ecology. For example, an inland population reaches the sea-coast or a savanna population enters the equatorial forest. In the most extreme case, an incoming population arrives from an exotic biogeographical zone of the world and encounters a fauna and flora that is almost entirely unfamiliar. This is the case in Madagascar, where Austronesians from island SE Asia, arriving around the 5th century AD, were faced with an largely endemic ensemble of plants and animals and were obliged to construct a new vocabulary to name them. Appendix 1. shows a sample of mammal names in Malagasy with my proposals for their etymologies. A very small proportion are directly from Austronesian, others derive from coastal Bantu and many are quite mysterious. Remarkably, even words for domestic animals such as 'cow' and 'pig' with which the Austronesians were familiar are replaced by Bantu. The likely explanation is that the Austronesians first reached, not Madagascar, but the East African coast. They enslaved local Bantu populations, who were carried to Madagascar to herd cattle and grow rice, and thus became more familiar with the endemic plants and animals than their masters (Blench 2010). A third element in Malagasy animal names are words with no clear provenance. It was previously thought that Madagascar was uninhabited prior to Austronesian settlement, but archaeological evidence suggests that low-density foragers from the African mainland reached the islands around 400 BC (Blench 2007b). Dewar et al. (2013) have recently claimed that stone tools push this date back to earlier than 2000 BC, although it is safe to say, that this is not yet widely accepted. There is presumably a link with the still-extant foragers, the Bēosi or Vazimba, whose language is now only recoverable as isolated lexicon from fragments recorded in the 1930s (Birkeli 1936; Blench & Walsh n.d.; Pierron et al. 2014). It is reasonable to suppose that animal names without etymologies in Malagasy are substrate elements from the prior languages of foragers who reached the island before the ancestors of the present-day occupants.

5.2 Fishing populations of the West African coast

The Ijò peoples, who today inhabit the Niger Delta in Nigeria, are presumed to have migrated from the Upper Niger several thousand years ago¹³. Some 3000 years ago there is a major increase in oil-palm pollen in the Delta (Şowunmi 1981) which probably points to an incoming population cutting down gallery forest. Whether the Ijò displaced an indigenous foraging population has not been resolved, but their ichthyological vocabulary is quite consistent, and yet shows no etymological affinities with either the names of river fish, nor signs of being constructed from language-internal resources. Appendix Table II gives a list of the principal fish names that can be reconstructed to proto-Ijò. Many vernacular terms for sea-fish names seem to appear from nowhere. This suggests strongly that there were pre-existing maritime populations specialised in catching pelagic species, and that the incoming Ijò adopted fish-names from them. Once these populations were assimilated, they persist in the vocabulary of sea-creatures.

5.3 Animal names of the ancient Sahara

Archaeological evidence points to an extremely long occupation of the Sahara (Drake, Blench et al. 2011), yet the languages spoken there today are clearly of recent origin. The Arabic-speaking Hassaniya that dominate Mauretania have all but assimilated the Zenaga, the last remaining Berber group (Taine-Cheikh 2010). The Central Sahara is virtually all Tuareg-speaking apart from the small island of Tetseret Berber in Niger (Lux 2011b). The Saharan languages, Teda-Daza and Beria [=Zaghawa], occupy the Chad-Sudan borderland and the eastern Sahara is again Arabic-speaking Bedawiin. Berber languages are extremely close to one another and cannot be very ancient. Blench (2001) identifies the spread of Berber languages with the east-west spread of ruminant livestock, principally cattle and sheep, which would put the Berber populations in Mauretania by about 4300 BP.

There are some populations in the Sahara which may be remnants of older foraging groups. The Imraguen, fishermen living along the sea-coast of Mauretania, the Nemadi migrating between Mali and eastern Mauretania and the Dawada, a now-dispersed people who subsisted on the endemic shrimps in the salt lakes of southern Libya. None of these peoples now speak a distinctive language, and their Arabic has not been systematically elicited to establish whether they retain specialised technical vocabulary which cannot be effectively etymologised.

Populations such as the Tuareg, moving from the oases of North Africa south into the desert and into the Sahel of West Africa would encounter unfamiliar plants and animals. They seem to have partly taken over names from the prior Berber populations, but some have no clear etymologies and they probably adapted these from the speech of forager populations the resident in the Sahara. Appendix III gives some examples of this lexicon.

¹³ One striking piece of evidence for this is the cognacy of the Ijò word for ‘manatee’ (*Trichechus senegalensis*) with the Mande terms, Proto-Ijò *imēĩ* and Bamana *mã*.

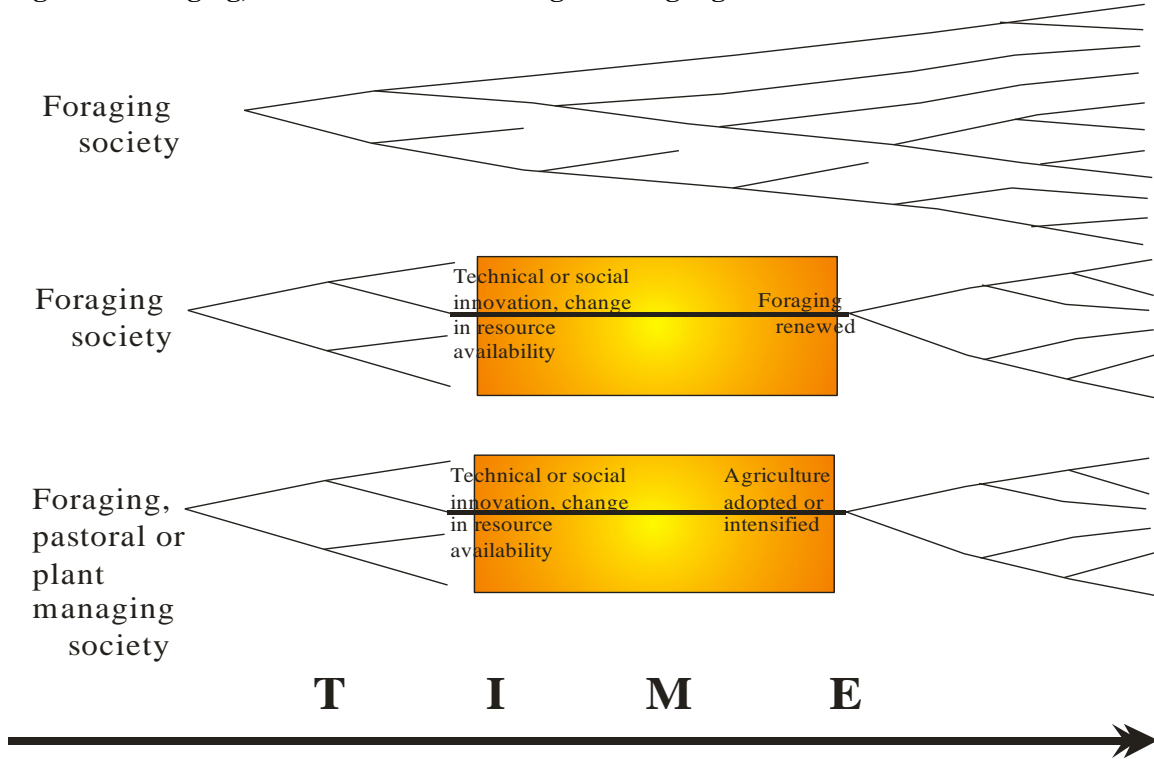
6. Conclusion: accounting for the linguistic situation in Africa

We cannot easily dispense with a stage in African prehistory when the continent was characterised by extreme linguistic and biological diversity. At the period when modern humans were diffusing from southern and eastern Africa, they would have spread over the continent at extremely low population densities either assimilating or out-competing existing hominid populations. Modern humans must have been interfertile with some resident African hominids, creating biological, social and linguistic complexity. Whether there was a pause before they entered West Africa remains a topic for further research. Most of these populations would have been physically unlike the Khoisanoids and resembled rather the Hadza, the Ongota, the Kwadi and the Damara (Blench 1999). If so, then this earlier diversity has largely disappeared and this must be accounted for by expansions of major language phyla. Such expansions are not unmotivated; there must be some economic or social innovation to account for them and the challenge is to trawl the archaeological and linguistic record for possible motivating factors. Ideally, the model would also explain the major differences between global language areas. The argument can be summarised as follows:

- a) Language goes back into the unknown past of foraging societies and probably to the genesis 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, and 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, the invention of the outrigger or the bow and arrow.
- c) Such changes provide a significant impetus to particular ethnolinguistic groups and they expand 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 distantiation, although there may be greater overall numbers of languages due to the increase 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.

Figure 2 presents these alternative processes as a diagram:

Figure 2. Foraging, the transition to farming and language diversification



Since the beginning of the Holocene there have been a series of phylic expansions within Africa, which have effectively eliminated its linguistic diversity. Fragmentary evidence for this remains in the few isolates still extant, as well as substrate lexicon. The analysis of substrate lexicon in Africa remains poorly developed, probably due to the intellectual tradition of ‘lumping’, although it is relatively easy to detect, especially in the lexicon of plants and animals.

Table 4 shows the phyla that now dominate Africa with my proposals for their origins, approximate dates and engines of dispersal (Blench 2006b). This proposal does not cover all of Khoisan, only the central Khoe languages, which are all closely related (Voßen 1996). The fragmentary nature of our documentation on other Khoisan branches may mean that these questions are in principle unanswerable.

Table 4. African language phyla: dates, homelands, archaeological correlations

Phylum	Date BP	Homeland	Engine	Correlation
Nilo-Saharan	>15,000	SW Ethiopia, Uganda	Climate improvements, fishing	'Green Sahara'
Niger-Congo	>10,000	Southern margins of the Sahara	New hunting techniques	Ounanian?
Afroasiatic	>10,000	SW Ethiopia	Livestock management	
Central Khoisan	>2000	South Central Africa	Livestock management	

The consequence has been that *bona fide* language isolates in Africa are very rare, and their status often debated. They must have been extremely common at some point, but have been assimilated and now can be detected only through substrate vocabulary. It seems likely that this is quite common, but the nature of scholarship applied to African languages has not been focused on this area. Ethnobiological research has declined drastically in recent years and modern dictionaries are often worse in this respect than those compiled half a century ago, probably because typical collaborators are semi-urbanised and now have no knowledge of the natural world. In addition, environmental degradation has meant that even those living in rural areas are unfamiliar with many species of plant and animal which would have been well-known a few decades ago. Just as linguists are much better at holding conferences deploring the disappearance of languages rather than going out and documenting them, so the same may also be true of research into the lexicon of the natural world and the detection of substrates.

As to the larger picture, we are still at the level of speculation. The situation in Africa may be said to resemble Eurasia, including SE Asia, where the expansion of a few phyla has almost completely eliminated diversity, leaving only scattered isolates and small phyla. This can be contrasted with Melanesia, Australia and the New World, where a greater diversity is the overall pattern, despite Holocene expansions of Pama-Nyungan and the Trans New Guinea phylum. It is noticeable that these three areas represent the 'end-of-the-line' as far as expansion of modern humans goes. They are geographically remote and thus were relatively isolated from technical and resource innovation, which drove language expansion. The situation in Africa and Eurasia seems to be that a large contiguous land area provides a breeding ground for innovation and crucially, its spread, thereby increasing the opportunities for innovation to be adopted and transform individual small forager societies. Hence the rather remarkable differences between regions of the world.

References

- Abbi, Anvita 2006. *Endangered Languages of the Andaman islands*. Munich: Lincom Europa.
- Abbi, Anvita 2012. *Dictionary of the Great Andamanese Language (English – Great Andamanese – Hindi)*. (First Edition with a CD-ROM). Delhi, India: Ratna Sagar.
- Ahland, C.A. 2010. Noun Incorporation and Predicate Classifiers in Gumuz. *Journal of African Languages and Linguistics*, 31(2):159-203.
- Anonby, Erik 2007. The labial flap in Mambay (Niger-Congo, Adamawa): Phonological rarity or fundamental element? *Journal of African languages and linguistics* 28:1, 1–17.
- Bar-Yosef, O. & Belfer-Cohen, A. 2001. From Africa to Eurasia-early dispersals. *Quaternary International*, 75: 19-28.
- Beekes, Robert S.P. 2010. The Pre-Greek Loanwords in Greek. In: *Etymological Dictionary of Greek*. Leiden: Brill.
- Behar, Doron M Richard Villems, Himla Soodyall, Jason Blue-Smith, Luisa Pereira, Ene Metspalu, Rosaria Scozzari, Heeran Makkan, Shay Tzur, David Comas, Jaume Bertranpetit, Lluís Quintana-Murci, Chris Tyler-Smith, R. Spencer Wells and Saharon Rosset 2008. The dawn of human matrilineal diversity. *American Journal of Human Genetics*, 82(5) : 1130-1140.
- Bender, M. L. 1975. *Omotic: A New Afroasiatic Language Family*. University Museum, Southern Illinois University, Carbondale.
- Bender, M. L. 1979. Gumuz: a sketch of grammar and lexicon. *Afrika und Übersee*, 62:38-69.
- Bender, M.L. 1969. Chance CVC correspondences in unrelated languages. *Language*, 45(3): 519-531.
- Bender, M.L. 1997. [2nd ed.] *The Nilo-Saharan languages: a comparative essay*. Munich: Lincom Europa.
- Bender, M.L. 2001. *Comparative Morphology of the Omotic Languages*. LINCOM Studies in African Linguistics 19. München: Lincom Europa.
- Bender, M.L. 2003. *Comparative Omotic Lexicon*. Carbondale: SIU.
- Bender, M.L. 1983. Remnant languages of Ethiopia and Sudan. in Bender, M.L. (ed.) *Nilo-Saharan Language Studies*. Michigan State University Press.
- Birkeli, Emil 1936. *Les Vazimba de la cote ouest de Madagascar : notes d'ethnologie*. Mémoires de l'Académie malgache. Tananarive : Imprimerie moderne de l'Emyrne, Pitot de la Beaujardière.
- Bleek, Dorothea 1956. *A Bushman dictionary*. New Haven: American Oriental Society.
- Bleek, Wilhelm H.I. 1862, 1869. *A comparative grammar of South African Languages*. (1862: Part I; 1869: Part II) London, Trübner & Co.
- Blench, Roger M. 1995. Is Niger-Congo simply a branch of Nilo-Saharan? In: *Proceedings of the Fifth Nilo-Saharan Linguistics Colloquium, Nice, 1992*. ed. R. Nicolai and F. Rottland. 68-118. Köln: Rudiger Köppe.
- Blench, Roger M. 2001. Types of language spread and their archaeological correlates: the example of Berber. *Origini*, XXIII: 169-190.
- Blench, Roger M. 1999. Are the African pygmies an ethnographic fiction? In: *Hunter-gatherers of Equatorial Africa*. K. Biesbrouyck, G. Rossel and S. Elders (eds.) 41-60. Leiden: Centre for Non-Western Studies.
- Blench, Roger M. 2002. Besprechungsartikel. The classification of Nilo-Saharan. *Afrika und Übersee*, 83:293-307.
- Blench, Roger M. 2006a. *The Kadu languages and their affiliation: between Nilo-Saharan, Niger-Congo and Afro-Asiatic*. Insights into Nilo-Saharan Language, History and

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

- Culture. Al-Amin Abu-Manga, L. Gilley & A. Storch eds. 101-127. Köln: Rüdiger Köppe.
- Blench, Roger M. 2006b. *Archaeology, Language and the African past*. Lanham: Altamira Press.
- Blench, Roger M. 2007a. Bangi Me: a language of unknown affiliation in Northern Mali. *Mother Tongue XII*: 147-178.
- Blench, Roger M. 2007b. New palaeogeographical evidence for the settlement of Madagascar. *Azania XLII*:69-82.
- Blench, Roger M. 2010. New evidence for the Austronesian impact on the East African coast. In: *Global origins and the development of seafaring*. Atholl Anderson ed. Cambridge: Macdonald Institute.
- Blench, Roger M. 2012. The role of agriculture in explaining the diversity of Amerindian languages. In: *The past ahead: language, culture and identity in the Neotropics*. Christian Isendahl ed. 13-37. Uppsala: Acta Universitatis Uppsaliensis.
- Blench, Roger M. 2013. Links between Cushitic, Omotic, Chadic and the position of Kujarge. In: *Proceedings of the 5th International Conference of Cushitic and Omotic languages*. Mari-Claude Simeone-Senelle & M. van Hove eds. 67-80. Köln: Rüdiger Köppe.
- Blench, Roger M. 2014. The origins of nominal affixes in MSEA languages: convergence, contact and some African parallels. In: *Mainland Southeast Asian Languages: The State of the Art 2014*. N.J. Enfield and Bernard Comrie eds. Mouton de Gruyter.
- Blench, Roger M. n.d. a. *Recovering data on Mpra [=Mpre] a possible language isolate in North-Central Ghana*. Electronic ms.
- Blench, Roger M. n.d. b. *A Dampo wordlist*. Electronic ms.
- Blench, Roger M. n.d. c. *Bere: a language of Côte d'Ivoire of unknown affiliation*. Electronic ms.
- Blench, Roger M. & M. Walsh n.d. *The vocabularies of Vazimba and Beosi: do they represent the languages of the pre-Austronesian populations of Madagascar?* Unpublished paper.
- Boyd, Raymond 1994. *Historical perspectives on Chamba Daka*. Köln: Rüdiger Köppe.
- Boyeldieu, P. n.d. *Présentation du láà:l ou "Gori" (Moyen-Chari, Tchad)*. ms. CNRS, Paris.
- Brown, Cecil n.d. *Statistical Demonstration of the Failure of Key Linguistic Evidence to Support the Amerind Hypothesis*. Circulated ms.
- Campbell, Lyle 1997a. *American Indian languages: the historical linguistics of native America*. New York/Oxford: Oxford University Press.
- Campbell, Lyle 1997b. Amerind personal pronouns: A second opinion. *Language* 73:339-351.
- Campbell, Lyle & W.J. Poser 2008. *Language classification*. Cambridge: Cambridge University Press.
- Cardinal, A.W. 1931. A survival. *Gold Coast Review*, 5(1):193-197.
- Cohen, M. 1947. *Essai comparatif sur le vocabulaire et la phonétique du Chamito-Sémitique*. Paris: Honoré Champion.
- Cohen, A.S., Stone, J.R., Beuning, K.R., Park, L.E., Reinthal, P.N., Dettman, D., Scholz, C.A., Johnson, T.C., King, J.W., and Talbot, M.R., et al. 2007. Ecological consequences of early Late Pleistocene megadroughts in tropical Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 104:16422-16427.
- Creissels, Denis n.d. *Liste de mots Pre*. ms.
- Crowley, T. and R.M.W. Dixon 1981. Tasmanian. In: *Handbook of Australian languages, Volume 2*. R.M.W. Dixon and B.J. Blake (eds). 394-427. Canberra: Australian National University Press.
- Dempwolff, Otto 1934. *Vergleichende Lautlehre des austronesischen Wortschatzes, Band 1: Induktiver Aufbau einer indonesischen Ursprache*. Beihefte zur Zeitschrift für Eingeborenen-Sprachen 15. Berlin: Dietrich Reimer.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

- Dempwolff, Otto 1937. *Vergleichende Lautlehre des austronesischen Wortschatzes, Band 2: Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen*. Beihefte zur Zeitschrift für Eingeborenen-Sprachen 17. Berlin: Dietrich Reimer.
- Dempwolff, Otto 1938. *Vergleichende Lautlehre des austronesischen Wortschatzes, Band 3: Austronesisches Wörterverzeichnis*. Beihefte zur Zeitschrift für Eingeborenen-Sprachen 19. Berlin: Dietrich Reimer.
- Dennell, R. and M. Petraglia 2012. The Dispersal of *Homo sapiens* across Southern Asia: How Early, How Often, How Complex? *Quaternary Science Reviews*, 47:15-22.
- Dewar, R.E., Radimilahy, C., Wright, H. T., Jacobs, Z., Kelly, G. O., & Berna, F. 2013. Stone tools and foraging in northern Madagascar challenge Holocene extinction models. *Proceedings of the National Academy of Sciences*, 110(31): 12583-12588.
- Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages*. Amsterdam: John Benjamins Publishing.
- Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Doke, C.M. 1961. The earliest records of Bantu. In: *Contributions to the history of Bantu linguistics*. C.M. Doke and D.T. Cole eds. 1-26. Johannesburg: Witwatersrand University Press.
- Doornbos, P. & M.L. Bender 1983. Languages of Wadai-Darfur. In: *Nilo-Saharan Language Studies*. M.L. Bender (ed.) 43-79. East Lansing: Michigan State University Press.
- Drake, N.A., Blench, Roger M., Armitage, S.J., Bristow, C.S., White K.H. 2011. Ancient watercourses and biogeography of the Sahara explain the peopling of the desert. *Proceedings of the National Academy of Sciences of America*, 108, 458-462.
- Ehret, Christopher 2001. *A historical-comparative reconstruction of Nilo-Saharan*. Köln: Rudiger Köppe.
- Fleming, Harold C. 1991. Shabo: presentation of data and preliminary classification. In: *Proceedings of the Fourth Nilo-Saharan Conference. Bayreuth, 1989*. Bender, M L. (ed.) 389-402. Hamburg: Buske.
- Fleming, Harold C. 2006. *Ongota: a decisive language in African prehistory*. Mainz: Harassowitz.
- Gilley, Leoma 2014. Katcha noun morphology. In: *New research in the languages of the Nuba mountains. Nuba mountains language studies*. Thilo Schadeberg & Roger Blench eds. 501-521. Köln: Rudiger Köppe.
- Greenberg, J.H. & Merritt Ruhlen 2007. *An Amerind etymological dictionary*. Department of Anthropological sciences, Stanford University.
- Greenberg, J.H. 1963. *The Languages of Africa*. The Hague: Mouton.
- Greenberg, J.H. 1971. The Indo-Pacific hypothesis. In: *Current Trends in Linguistics, Volume VIII - Linguistics in Oceania*, T.A. Sebeok (ed.) 807-871. The Hague: Mouton.
- Greenberg, J.H. 1987. *Language in the Americas*. Stanford, CA: Stanford University Press.
- Greenberg, J.H. 1996. In defense of Amerind. *International journal of American linguistics*, 62(2): 131-164.
- Greenberg, J.H. 2000. *Indo-European and its closest relatives: the Eurasiatic language family*. Stanford, CA: Stanford University Press.
- Grün, R., Stringer, C., McDermott, F., Nathan, R., Porat, N., Robertson, S. et al. 2005. U-series and ESR analyses of bones and teeth relating to the human burials from Skhul. *Journal of Human Evolution*, 49: 316-34.
- Güldemann, Tom 2004. Reconstruction through 'de-construction': the marking of person, gender and number in the Khoe family and Kwadi. *Diachronica*, 21(2):251-306.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

- Güldemann, Tom 2008. The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In: Bernd Heine & Derek Nurse (eds), *A linguistic geography of Africa*. 151-185. Cambridge: Cambridge University Press.
- Güldemann, Tom 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In: *Geographical typology and Linguistic areas*. O. Hieda, C. König and H. Nakagawa eds. 109-142. Amsterdam/Philadelphia: John Benjamins.
- Gunz, P., Bookstein, F.L., Mitteroecker, P., Stadlmayr, A., Seidler, H. & Weber, G.W. 2009. Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario. *Proceedings of the National Academy of Sciences of America*, 106:6094-6098.
- Hall, Ed and Marian Hall 2004. Kadugli-Krongo. *Occasional Papers in the Study of Sudanese Languages*, No. 9. Khartoum.
- Hammarström, H. 2012. Pronouns and the (Preliminary) Classification of Papuan languages. *Language and linguistics in Melanesia*, Special issue 2012, 2: 428-539.
- Heath, Jeffrey 2006. *Dictionnaire touareg du Mali*. Paris: Karthala.
- Heine, Bernd & Derek Nurse (eds.) 2008. *A linguistic geography of Africa*. Cambridge: Cambridge University Press.
- Heubeck, Alfred 1961. *Praegraeca: sprachliche Untersuchungen zum vorgriechisch-indogermanischen Substrat*. Universitätsbund Erlangen.
- Hyman, L.M. 2011. The Macro-Sudan Belt and Niger-Congo Reconstruction. *Language Dynamics and Change*, 1(1):3-49.
- Jackson, K. 1955. The Pictish Language. In: Wainwright, F.T. ed. *The Problem of the Picts*. 129-166. Edinburgh: Nelson.
- Ke, Yuehai *et al.* 2001. African Origin of Modern Humans in East Asia: A Tale of 12,000 Y Chromosomes. *Science*, 292: 1151-1153.
- Kleinewillinghöfer, U. 2001. The language of the Jalaa: a disappearing language isolate. *Sprache und Geschichte in Afrika*, 16/17: 239-271.
- Lewis, P. *et al.* 2013. *Ethnologue Language Family Index (17th edition)*, Dallas: Summer Institute of Linguistics.
- Lionnet, F. (n.d.) *Laal: an isolate language?* Powerpoint presented in Lyon, November 2010.
- Lockwood, W.B. 1975. *Languages of the British Isles, Past and Present*. London: André Deutsch.
- Lovestrand, J. 2012. *Classification and description of the Chadic languages of the Guéra (East Chadic B)*. SIL Electronic Working Papers 2012-004. Dallas, Texas: SIL International.
- Lux, Cécile 2011. *Étude descriptive et comparative d'une langue menacée : le tetseret, langue berbère du Niger*. Ph.D. Université Lumière Lyon 2 École doctorale : Lettres, langues, linguistique, arts.
- Macaulay, Vincent *et al.* 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*, 308:1034-1036.
- MacDougal, Ian, Francis H Brown, & John G. Fleagle 2005. Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature*, 433: 733-736.
- Masica, C.P. 1993. *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
- Mfonyam, Joseph 1989. *Tone in Orthography: The Case of Bafut and Related Languages*. PhD Thesis. University of Yaoundé 1.
- Olson, Kenneth S. and John Hajek 2003. Crosslinguistic insights on the labial flap. *Linguistic Typology*, 7,2: 157-186.
- Painter, C. 1967. The distribution of Guang in Ghana and a statistical pre-testing on twenty-five idiolects. *Journal of West African Languages*, 4,1:25-78.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

- Pawley, Andrew 2005. The chequered career of the Trans New Guinea hypothesis: recent research and its implications. In: *Papuan Pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*. Andrew Pawley, Robert Attenborough, Jack Golson and Robin Hide (eds.) 67-108. PL 572. Canberra: ANU.
- Pawley, Andrew 2012. How reconstructable is proto Trans New Guinea? Problems, progress, prospects. *Language and linguistics in Melanesia*, Special issue 2012, 1: 88-164.
- Petraglia, M.D. 2007. Mind the gap: factoring the Arabian peninsula and the Indian subcontinent into Out of Africa models. In: *Rethinking the Human Revolution*. P. Mellars, O. Bar-Yosef, K. Boyle & C. Stringer eds. 383-394. Cambridge: McDonald Institute Archaeological Publications.
- Petraglia, M.D. and J.I. Rose (eds.) 2009. *The Evolution of Human Populations in Arabia*. Dordrecht: Springer Verlag.
- Petraglia, M.D., Haslam, M., Fuller, D.Q. & Boivin, N. 2010. The southern dispersal route and the spread of modern humans along the Indian Ocean rim: New hypotheses and evidence, *Annals of Human Biology*, 37(3): 288-311.
- Pierron, Denis, Harilanto Razafindrazaka, Luca Pagani, François-Xavier Ricaut, Tiago Antao, Mélanie Capredon, Clément Sambo, Chantal Radimilahy, Jean-Aimé Rakotoarisoa, Roger M. Blench, Thierry Letellier, Toomas Kivisild 2014. Genome-wide evidence of Austronesian–Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. *Proceedings of the National Academy of Sciences of the United States of America*, <http://www.pnas.org/content/early/2014/01/03/1321860111.abstract>
- Rankin, Robert L. 2012. Deeper Genetic Relationships in North America: Some Tempered Pessimism. In: *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 22(2): 117-128.
- Rankin, Robert L. 1992. Review: Joseph H. Greenberg, Language in the Americas. *International Journal of American Linguistics*, 58:324-351
- Ringe, Don A. 1992. [On calculating the factor of chance in language comparison](#). *Transactions of the American Philosophical Society*, 82(1):1-110.
- Ringe, Don A. 1999. Language classification: scientific and unscientific methods. In: *The human inheritance: genes, language, and evolution*. Brian D. Sykes ed. 45-74. Oxford: Oxford University Press.
- Rose, J. I., Usik, V. I., Marks, A. E., Hilbert, Y. H., Galletti, C. S., Parton, A., ... & Roberts, R. G. 2011. The Nubian complex of Dhofar, Oman: an African middle stone age industry in southern Arabia. *PLoS 1*, 6(11): e28239.
- Ruhlen, Merritt 1994. Amerind T'A?NA 'Child, Sibling. In: *On the Origin of Languages*. M. Ruhlen 183–206. Stanford University Press.
- Ruhlen, Merritt 1995. Proto-Amerind Numerals. *Anthropological Science* (Tokyo) 103:209–25.
- Sands, B. 1998. *Eastern and Southern African Khoisan: evaluating claims of a distant linguistic relationships*. Quellen zur Khoisan-Forschung 14. Köln: Rüdiger Köppe.
- Savà, Graziano and Mauro Tosco 2000. A sketch of Ongota: a dying language of Southwest Ethiopia. *Studies in African Linguistics*, 29:2: 59-135.
- Sidwell, Paul & R.M. Blench 2011. The Austroasiatic *Urheimat* : the Southeastern Riverine Hypothesis. In: N. Enfield ed. *Dynamics of Human Diversity in Mainland SE Asia*. Canberra: Pacific Linguistics.
- Sommer, Gabriele and Rainer Voßen 1992. Schnalzwörter im Yei (R.41). *Afrika und Übersee*, 75:1-42.
- Şowunmi, M.A. 1981. Late Quaternary Environmental Changes in Nigeria. *Pollen et Spores*, 13(1):125-148.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

- Stevenson, Roland C. 1991. Relationship of Kadugli-Krongo to Nilo-Saharan: Morphology and lexis. In: *Proceedings of the Fourth Nilo-Saharan Conference*. Bayreuth, 1989. Bender, M L. (ed.) 347-369. Hamburg: Buske.
- Storch, Anne 2005. *The Noun Morphology of Western Nilotic*. Köln: Köppe.
- Sudlow, David 2009. *Dictionary of the Tamasheq of North-East Burkina Faso*. Köln: Rudiger Köppe.
- Sweetman, Gary 1981a. *Lexique francais-fali-fulfulde*. ms. Yaoundé: SIL.
- Sweetman, Gary 1981b. *A comparative stud of fali dialects*. ms. Yaoundé: SIL.
- Taine-Cheikh, C. 2010. *Dictionnaire français-zénaga (berbère de Maurétanie)*. Köln: Rudiger Köppe.
- Trombetti, Alfredo 1905. *L'unità d'origine del linguaggio*. Bologna: Libreria Treves di Luigi Beltrami.
- Van Peer, P. 1998. The Nile Corridor and the Out-of-Africa Model: An Examination of the Archaeological Record. *Current Anthropology*, 39:115-140.
- Vermeersch, P.M. 2001. Out of Africa' from an Egyptian point of view. *Quaternaire International* 75: 103-112.
- Voßen, R. 1996. *Die Khoe-sprachen*. Köln: Rudiger Köppe.
- Westermann, Diedrich 1911. *Die Sudansprachen*. Hamburg: Friederichsen.
- Westermann, Diedrich 1927. *Die westlichen Sudansprachen und ihre Beziehungen zum Bantu*. Berlin: de Gruyter.
- Westermann, Diedrich 1935. Nominalklassen in westafrikanischen Klassensprachen und in Bantusprachen. *MSOS*, 38: 1-52.
- Westphal, E.O.J. 1963. The linguistic prehistory of Southern Africa: Bush, Kwadi, Hottentot and Bantu linguistic relationships. *Africa*, 33: 237-265.
- Whitehouse, P., Usher, T., Ruhlen, M., & Wang, W.S.Y. 2004. Kusunda: An Indo-Pacific language in Nepal. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15): 5692-5695.
- Wilson, John G. 1970. Preliminary observations on the Oropom people of Karamoja, their ethnic status, culture and postulated relation to the peoples of the late stone age. *Uganda Journal*, 34(2):125-145.

Appendices

I. Malagasy mammal names

No.	Malagasy	English	Scientific name	Etymological commentary
1.	amboa laolo	Falanouc	<i>Eupleres goudotii</i>	cf. amboahaolo 'feral dog' < amboa 'dog' haolo 'wild' [R]. amboa is from a Bantu source, e.g. Comorian & Swahili mbwa (9/10) 'dog' < Proto-Sabaki *(i)mbwa (9/10) 'dog' [N&H]. The transfer from 'dog' to 'falanouc' (an endemic carnivore) is perhaps a secondary derivation in Malagasy
2.	ampongy	Eastern avahi	<i>Avahi laniger</i>	cf. Swahili (Unguja) k^hima punju (9/10) 'Zanzibar red colobus, <i>Colobus kirkii</i> '; also Nyakyusa kipunji (7/8) 'Highland mangabey, <i>Rungwecebus kipunji</i> ' [Davenport et al.]. The Malagasy term for this medium-sized lemur may be derived from a form of the Bantu monkey name with class 3 prefix (?* mpungi)
3.	andrehy [G&B]	Fruit bats	Pteropodidae	Richardson (1885: 43) defines this as the name of a bird. cf. Comorian (Ndzuan) ndrege (9/10) 'bird (generic)'; Swahili ndege (9/10) 'bird'. This is an innovation in the southern dialects of Swahili, probably borrowed from one of the mainland Bantu languages [N&H]. In Swahili and related languages bats are often classified as birds
4.	ankomba, komba	Crowned lemur	<i>Eulemur coronatus</i>	cf. Swahili (Unguja) k^homba (9/10) 'galago spp.' [Walsh] < Proto-Sabaki *nkomba (9/10) 'galago' [N&H]. Given the resemblance between these two groups of primates, the transfer of a name from galagos (= bushbabies) to lemurs, which are indigenous to Madagascar, is as natural as was the former English practice of referring to bushbabies as 'lemurs'
5.	antsanga	Bushpig	<i>Potamochoerus larvatus</i>	cf. Swahili (Unguja) kitanga (7/8) 'solitary male bushpig'? The Malagasy form is possibly derived from an earlier ?*ncanga (9/10) 'male bushpig' < Proto-Sabaki *-canga v 'to wander' [N&H]

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

No.	Malagasy	English	Scientific name	Etymological commentary
6.	antsangy	rice tenrecs	<i>Oryzorictes spp.</i>	cf. Swahili (Tanzanian mainland) sange (9/10) 'elephant shrew spp.' [Swynnerton]; Mijikenda (Giryama) tsanje (?ts ^h anje) (9/10) 'Four-toed elephant shrew, <i>Petrodomus tetradactylus</i> ' [Costich] < earlier ?* ntsange . This term has widespread cognates in Tanzania. Elephant shrews are superficially similar to the endemic rice tenrecs of Madagascar
7.	boenga, boengy	Milne-Edward's sportive lemur	<i>Lepilemur edwardsi</i>	cf. Sungai (East Sabah) bongan 'Hose's langur'
8.	falanouc	Falanouc	<i>Eupleres goudotii</i>	cf. Barito lects also Lun Dayeh (Sabah) pəlanuk 'mouse-deer spp.' (<i>Tragulus napu</i> , <i>T. javanicus</i>). A strange semantic shift but the form is very close. However, both the size and posture of these two species are not dissimilar
9.	fanihy	Fruit bats	Pteropodidae	cf. PMP * paniki 'flying fox'. Blust (2002: 107) notes that reflexes of this are <i>absent</i> in Borneo and thus the reflex in Malagasy is rather surprising [see Adelaar on other sources for Malagasy]
10.	fosa	Fossa	<i>Cryptoprocta ferox</i>	Beaujard derives this from purported < PMP 'cat' but as Blust (2002: 99) points out, Western Austronesian forms such as Iban posa are almost certainly derived from poes and these are convergent borrowings from the 17 th century. The Malagasy term may therefore be a late and independent borrowing from a trade language. However, it turns out that pusa and similar are also Malay for the Malay weasel (<i>Mustela nudipes</i>) which may have an old anthropic distribution in the region. It therefore may have shifted to 'cat' in island SE Asia and to fossa in Madagascar. Cf. Malagasy bosy 'feral cat' [R] < Swahili busi [not in dictionaries] < Arabic [Simon]; also Nyakyusa pusi (1a/2) 'cat' [Felberg]
11.	gidro	Crowned lemur	<i>Eulemur coronatus</i>	Richardson (1885: 213) suggests a comparison with Swahili ngedere

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

No.	Malagasy	English	Scientific name	Etymological commentary
				'monkey sp.' and/or Arabic qird 'ape'. Simon (1988: 291) supports an "araboswahili" etymology. ngedere (9/10) is a southern Swahili dialect name for the Blue monkey, <i>Cercopithecus mitis</i> , presumed to be borrowed from a neighbouring Bantu language [N&H]. If corroborated this would represent another example of a monkey name transferred to a lemur
12.	kelora	Common tenrec	<i>Tenrec ecaudatus</i>	? PCEMP * kandoRa 'cuscus, phalanger', e.g. Watubela kadola . cf. Blust (2002: 110) though these are east of the usual sources of Malagasy
13.	lambo	Bush pig	<i>Potamochoerus larvatus</i>	< Malay lambu , bovine, the original meaning, surviving in special expressions [Bj]
14.	radjako rajako [Simon] jakoe, jakoey [Gn]	Perrier's sifaka	<i>Propithecus diadema perrieri</i>	< French jacquot (<i>faire le jacques, faire le singe</i>), an abusive term for lemurs [Simon]. Also Indian Ocean Creole zako ; Comorian djakwe [Gn]
15.	tandraka, trandraka tandeke [R]	Common tenrec	<i>Tenrec ecaudatus</i>	? cf. Malay landak 'porcupine' (Adelaar 1989) and secondary borrowing into Comorian Ngazidja landa Ndzuwani landrá , Maore landra (9/10) 'tenrec'
16.	tranga lavaka	Small-toothed sportive lemur	<i>Lepilemur microdon</i>	cf. Kadazan (Sabah) tangah 'flying lemur'.
17.	varika	lemur spp.		< Maanyan warik 'monkey sp.' [Bj < Dahl]
18.	voalavo, valàvo valave [R]	rat spp.		cf. Proto-Austronesian * labaw , e.g. Kayan lavo , Muna (Sulawesi) wulawo . N.B. there are apparently no reflexes of the common SE Borneo * lésu (Blust 2002: 107).
19.	vontsira	Ring-tailed mongoose	<i>Galidia elegans</i>	cf. Swahili (Unguja) (9/10) nguchiro 'Banded mongoose, <i>Mungos mungo</i> ' (an introduced sp.) [Pakenham]; probably a loanword from a Rufiji-Ruvuma language, cf. Ngindo lingwichiro 'Banded mongoose, <i>Mungos mungo</i> '

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

No.	Malagasy	English	Scientific name	Etymological commentary
				[Stronach et al.]; Matumbi ngwicho 'mongoose spp.' [Stronach et al.]; also Pogoro lingwichiro 'Dwarf mongoose, <i>Helogale parvula</i> , & Banded mongoose, <i>Mungos mungo</i> ' [Stronach et al.]

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

1. II. Proto-Ijo fish names without discernible etymologies

English	Latin	proto-Ijo
Angelfish	<i>Pomacanthus paru</i>	ofún
Barracuda	<i>Sphyraena spp.</i>	ḍoṛo
Cameroon Sardine, Shad		bala
Cassava Croaker	<i>Pseudotolithus senegalensis</i>	oṗúla
Catfish	<i>Arius spp.</i>	ṣìngì, unction
Conger Eel	? <i>Paraconger notialis</i>	ábònyon
Flying Fish	<i>Cheilopogon sp.</i> ,	mindì-ófonì
Globefish	<i>Ephippion guttifer</i>	ububu
Grunt, burro	<i>Plectorhynchus spp. & Parapristipoma spp.</i>	eḡelèu
Hammerhead Shark	<i>Sphyrna spp.</i>	ugubéri
Horse Mackerel	<i>Caranx spp.</i>	kidoghó
Ladyfish	<i>Elops senegalensis</i>	bala-dowoin
Little Sleeper	<i>Eleotridae</i>	kalá ukulì
Mangrove Sleeper	<i>Eleotridae sp.</i>	ikúli
Mudskipper	<i>Periophthalmus spp.</i>	atílaì
Pigfish	? <i>Pomadasyis rogerii</i>	osisi
Ribbonfish	<i>Trachypterus trachypterus</i>	bini ópiya, gbógó
Royal Threadfin	<i>Pentanemus quinquarius</i>	inda
Sawfish	<i>Pristis spp.</i>	oki
Slender Long-Nosed Shark	? <i>Carcharinus signatus</i>	ofuruma
Snapper	<i>Lutjanus spp.</i>	agbará, tòmì
Soapfish	<i>Rypticus saponaceus</i>	erému
Sole		bumopalì
Spanish Mackerel	<i>Scomberomorus tritor</i>	sonoma-siko
Spotted Eagle-Ray ?	<i>Aetobatus narinari</i>	eḡbèin
Sting Ray	<i>Dasyatis spp.</i>	ṣìka
Tarpon	<i>Megalops atlanticus</i>	imunú dowoi
Tilapia		tabalá
Toadfish	<i>Batrachoididae spp.</i>	bini óḡumu
Turbot		saídu
Weakfish	<i>Atractoscion aequidens</i>	ona
Jellyfish		álapá
Common Periwinkle	<i>Tympanotonus fuscatus</i>	iṣembí
Dogwhelk	<i>Nucella lapillus</i>	igbekete
Oyster Spat	<i>Ostrea spp.</i>	imgbe
Cuttlefish, Squid		burumizi

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

III. Large animals names in Southern Berber languages

Many of the animal species the Berber encountered as they expanded south into Sub-Saharan Africa were unfamiliar in the Maghreb. The names for these are extracted from the relevant dictionaries. Very few have any obvious etymologies and the assumption is that these were adapted from now-vanished forager populations.

Tetserret animal names

Table 5. Animal names in Tetserret of Niger

English	French	sg.	pl.	Comment
antelope	antilope	ænar ^f		? Dorcas gazelle. Common South Berber
cat, wild	chat sauvage	mzuru		? < Hausa
chameleon	caméléon	tawit		
crow	corbeau		ɔyɾut-ən	
gazelle	gazelle	aʒonkəɖ		
giraffe	girafe	(ə)ʃiyeg	ʃiyeg-ən	
hare	lapin, lièvre	tmarwult, t-əmarwəl-t		
hawk	épervier, aigle	ənollam		
hedgehog	hérisson	tarangat		
hyena	hyène	tafagant		also the name for someone with strange habits (without female markers)
jackal	chacal	eridel		
lion	lion	ar	arr-ən	cf. Zenaga
ostrich	autruche	arəg	argan	
sheep, wild	mouflon	arig	argan	
vulture	vautour, charognard	abəngadew		

Zenaga animal names in Taine-Cheikh (2010)

Table 6. Animal names in Zenaga of Mauretania

English	French	sg.	pl.	Hassaniya
aardvark	oryctérope	təkši-n-tutfən		' <i>ovin-caprin des fourmis</i> '
bustard	outarde	ägäyš	əgäyššän	
crow	corbeau	täyyäL	täyyäyən	
dama gazelle	gazelle dama	änaʔr	änaʔrän	cf. Tamachek
dorcas gazelle	gazelle dorcas	äžänkud	äžänkud	<i>Gazella dorcas</i> . ġzāl
eagle	aigle	aʔdʔi	āʔdʔün	
elephant	éléphant	iyih	iyān	
francolin	perdrix	tuffurdah	tfurdäyn	

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

English	French	sg.	pl.	Hassaniya
gazelle, red-fronted	gazelle à front rouge	ažəmmi	ižəmma'n	<i>Gazella rufifrons</i> dāmi
giraffe	girafe	ažraf	əžraffän	cf. Arabic
hare I	lièvre I	tärämbuL	trumbäyən	
hare II	lièvre II	tyəžzaZ		
hawk	épervier	tumdah	tumdayn	
hedgehog	hérisson	gänvud	əgnävīd	? < H.
hyena, spotted	hyène, tacheté	gähūh		<i>Crocuta crocuta</i> . ḍab ^s
hyena, striped	hyène, rayé	ärḍäy	ərḍäyän	
jackal	chacal	äžḍih	äžḍādän	
leopard	panthère	ağayniš	əğayniššän	
lion	lion	wa'r	wa'rän	cf. Tetseret
monitor lizard	varan	kuḍih		
monkey	singe	äbugäy	əbugäyän	
oryx	oryx	wərg		< H.
pelican	pélican	ädänāy		< H.
pigeon	pigeon	i'milli	a'mällän	
python	python	girižmä		? < H.
rat	rat	əmiḍniš	əməḍnäššän	
ratel	ratel	ämässäf ən üržän		'dechireur de tendons d'Achille'
roan antelope	hippotrague	äžämiy	izämäyän	äžäməl. cf. Mali Tamachek
sheep, wild	mouflon	äräwih	əräwän	
viper	vipère	täššuffäh	täššuffäyän	
vulture	vautour, 'aigle'	ägo'dər	əgo'dərän	cf. Tamachek
warthog	phacochère	a'ž(ž)iy-ən	tnäyri ^h	'âne de la brousse'

Heath's Mali Tamachek dictionary (2006)

Table 7. Animal names in Tamachek of Mali

English	French	sg.	pl.	Comment
aardvark	oryctérope	taläwləwt		
bat	chauve-souris	a-færtætṭa		
buffalo	buffle	ésù		usual word for 'cow'
bustard	outarde	a-s-ækkætṭay		'one who shows his buttocks'
cane rat I	aulacode I	æ-kündär		but also mouse sp., also dassie
cane rat II	aulacode II	t-e-məlu-t-t		but also mouse sp., also dassie
chameleon	caméléon	t-ä-haw-t		
cheetah	guépard	á-dhal		
crocodile	crocodile	zəngəwáy		also water monitor lizard
crocodile	crocodile	ə-yata		also land monitor lizard
crow	corbeau	tàkrit		
elephant	éléphant	élu	älwanən	

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

English	French	sg.	pl.	Comment
fennec fox	renard	e-žægæž		<i>Fennecus</i> (= <i>Vulpes</i>) zerda;
fennec fox	renard	æ-kórsi		
fox, general	renard	e-šájæš		
francolin	perdrix	æ-fukáret		
gazelle, dama	gazelle dama	t-è-nher-t		
gazelle, dorcas	gazelle dorcas	t-ašňkətt		<i>Gazella dorcas</i>
gazelle, red- fronted	gazelle à front rouge	e-dám	i-dáman	<i>Gazella rufifrons</i>
genet	genet	éltæγ		
giraffe	girafe	á-mdəy		
hare	lièvre	t-e-mærwæl-t		
hedgehog	hérisson	t-e-kænəsi-t-t		
hippo	hippopotame	æ-jamba		< Songhay
honey	ratel	a-fæzæza		
badger				
hooded	vautour	borkíya		
vulture				
hyena, spotted	hyène, tacheté	šæbójæn		<i>Crocota crocuta</i>
hyena, striped	hyène, rayé	a-rídal		
jackal	chacal	e-bægg	ì-bæggan	<i>Canis adustus?</i> also <i>C. aureus</i>
jerboa	jerboa	e-ǧæww		<i>Jaculus jaculus</i> , also gerbil
leopard	panthère	wášil		
lion	lion I	á-hærr		
lion	lion II	a-wæqqas		
monkey	singe	ə-bíddæw		
monkey, patas	singe	kæyá		
oryx	oryx	óræj		<i>Oryx dammah</i> , now perhaps extinct locally
oryx	oryx	t-i n ísæk		<i>Oryx dammah</i> , now perhaps extinct locally
oryx	oryx, topi	t-e-dæri-t-t		
ostrich	autruche	é-tæqq		
ostrich	autruche	a-náhil		
ostrich	autruche	góyba		
pelican	pélican	t-oræf-t n am-an		'x of water'
pigeon	pigeon I	e-dæber		
pigeon	pigeon II	kálláya		<i>Oena capensis</i>
pigeon	pigeon III	t-əzún		
speckled predator	prédateur	bèr-wæqqas		

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

English	French	sg.	pl.	Comment
python	python I	t-a-žæbdær-æt		perhaps <i>Gongylophis</i> (= <i>Eryx</i>) <i>muelleri</i> cf. Songhay <i>namey hasu</i>
python, rock	python II	t-ànæyw-æt		<i>Python sebae</i>
roan antelope	hippotrague	a-šæmol		perhaps the roan antelope <i>Hippotragus equinus</i> , now locally extinct like other antelopes. cf. Zenaga
viper	vipère	kætètunḡu		<i>Cerastes cerastes</i> Songhay phrase “bring a wrap (= shroud)!”]
viper	vipère	s-æffæltæs		<i>Cerastes vipera</i> but also spitting cobra
viper, spotted	vipère	t-e-bæki-t-t		also horned <i>Cerastes</i> <i>cerastes</i>
vulture	vautour	e-žædær		<i>Gyps rueppellii</i>
vulture	vautour	t-æ-ɣalje		<i>Neophron percnopterus</i> ,
vulture	vautour	e-jædær, e-žædær		<i>Gyps rueppellii</i>
warthog I	phacochère I	a-gæŋgæra		
warthog II	phacochère II	a-zæybæra		
zorilla	zorille	a-ræraŋḡa		<i>Ictonyx striatus</i>

Sudlow's Burkina Faso Tamachek dictionary (2009)

Table 8. Animal names in Tamachek of Burkina Faso

English	French	sg.	Comment
aardvark	oryctérope	adhæg	
addax	addax	aməllal	
antelope	antilope	əzām	
antelope sp.	antilope sp.	tasārakənt	
baboon I	babouin I	kāya	
baboon II	babouin II	abiddāw	
bat	chauve-souris	afārtātta	
buffalo	buffle	esu n āroḡḡ	
bustard	outarde	āgayəs	
chameleon	caméléon	tawət	
cheetah	guépard	adal	
civet	civette	telḡət	
cobra, spitting	cobra cracheur	āssām (Z)	
cobra, spitting	cobra cracheur	emāḡāl (S)	
crane	grue	tenek	
crocodile	crocodile	āḡata	
dama gazelle	gazelle dama	ener	cf. Zenaga
dorcas gazelle		ažənkəḡ	<i>Gazella dorcas</i>

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory • Issue XVIII • 2013
50th Anniversary of J.H. Greenberg's *The Languages of Africa* (1963)

English	French	sg.	Comment
elephant	éléphant	elu	
fennec fox	renard	ezāgāz	
francolin	francolin	āfukāret	
gazelle sp.	gazelle sp	abugāra	
gazelle, red-fronted	gazelle à front rouge	edāmi	<i>Gazella rufifrons</i>
gerbil		eḏōwi	
giraffe	girafe	amdāy	
ground squirrel	écureuil de terre	ākolān	
guinea fowl	pintade	taylālt	
hare	lièvre	tekyāryālt	
hedgehog	hérisson	tekōneššit	
hippo	hippopotame	āgamba	< Songhay
hyena, spotted	hyène, tachetée	tāžori	<i>Crocuta crocuta</i>
hyena, striped	hyène, rayé	aridal	
jackal	chacal	ebāggi	<i>Canis adustus</i>
jackal	chacal	in-tānyōn	
leopard	panthère	wašil	
lion I	lion I	āhārr	
lion II	lion II	āxxu	
lion III	lion III	ebāyāw	
monitor lizard	varan	āyata	
monkey	singe	abbidāw	patas monkey
monkey	singe	kāya	
mouse	souris	immi	
oryx	oryx	ezām	
ostrich	autruche	anil	
ostrich	autruche	etāqq	Mali dialect
ostrich male	autruche, male	abal	
python	python	tanāywāt	
rat	rat	ākundār	
ratel	ratel	afāzāza	
rhino	rhino	tin isək	
roan antelope	hippotrague	ažāmōl	
topi	topi	edəri	
viper	vipère	taššelt	
vulture	vautour	egādār	cf. Zenaga
vulture	vautour	əziz	
vulture	vautour	tāyalge	
warthog I	phacochère I	agāṅgāra	
warthog II	phacochère II	ažāybāra, azebāra	
zorilla	zorille	fəkədərri	<i>Ictonyx striatus</i>
zorilla	zorille	gāṅgāhila	<i>Ictonyx striatus</i>
zorilla	zorille	arārāṅṅa	<i>Ictonyx striatus</i> Mali