

# How can Evidence from Current Day Languages be used to Illuminate our Understanding of Early Human Evolution and Dispersal?

Jahan ‘Harry’ Rezakhanlou-Taubman \*

## 1 Introduction

This question deals with the interplay between the analysis of the development of language on the one hand, and two distinct areas of analysis on the other: the evolution of our species, and the migration patterns of communities. It presents the challenge of integrating linguistic evidence with genetic and archeological data to support findings that go well beyond the scope of linguistics as such, and using linguistic evidence to illuminate our understanding of the wider phenomena of human evolution and the historical dispersal of humanity across the globe. Each of these questions in isolation raise fundamental challenges in their separate disciplines of evolutionary biology and human geography, so it is difficult to overestimate the challenge of attempting convergence between these three areas. At the same time, if the direct evidence presented by genetics and archeological sites may support analysis of the historical development of language, particularly concerning early language use for which there is inevitably no written evidence, it strengthens the respective claims made by linguists.

Nevertheless, the question of how language itself emerged remains inherently difficult, given the lack of direct evidence available and the difficulty of using the little available evidence to trace the nature of the remote origins of language. This is especially the case if one is using ‘current day’ language, since it means first developing a theory of how language evolves, and then in effect reversing time and using that theory to follow linguistic evolution from today’s usage patterns all the way back to the source of language. The complexity increases when this line of research is to be coordinated with research on evolution and human dispersal.

Some have in fact explained the development of human language in terms of evolution, for instance Pagel (in Wade, 2011, commenting on Atkinson, 2011) pointed out that “language was our secret weapon, as soon as we got language we became a really dangerous species,”. As an enormous factor driving the direction of our evolution as a species, it is logical to assume that language and the human species evolved together through everyday social interaction and necessity (Schoenemann, 2009). Considering the link between evolution and language development, Tagliatela et al., 2008, p. 343 conclude that “the neurological substrates underlying language production in the human brain may have been present in the common ancestor of

---

\*This paper was originally submitted for the subject Exploring Linguistic Diversity (LING30001) in 2016

humans and chimpanzees”, and that these sections of the brain in chimpanzees are used in communication functions. Language as such is not unique to *homo sapiens* – or even to the genus *homo* – and this makes it difficult to fix a definite period for the ‘origin of language’ – the question of determining an ‘origin’ even raises questions about the boundaries of what we consider to be ‘language.’ Recent genetic research (Penisi 2013) has disclosed interbreeding between species, so that the genetic lineage of some *homo sapiens* today includes Neanderthal and Denisovan genes. The fact that there is archeological evidence that proves these species cohabited some locations gives rise to speculation that at least some forms of language may have been shared between species. A notable example was the Neanderthal hyoid bone – a horseshoe-shaped bone in the neck above the larynx supporting the base of the tongue, essential for the production of the sounds used in human speech (as the jaw, tongue and larynx muscles are anchored to it) – found in Mugharat al-Kabara, Occupied Palestine, showing that the physical ability for language amongst protohumans dates back at least 60,000 years (Valadas et al., 1987).

## 2 Using Contemporary Language Patterns to Trace the Past Development of Language

Thus, while conceptually distinct from the study of human evolution, linguistic analysis is intertwined with the strictly biological aspects of the study and tracing dispersal as a matter of human geography. The essential question then is how, and whether, we can learn about historic development of language (‘diachronic’ in the term introduced by Saussure) through an investigation of aspects of contemporary language use, such as phonetics or syntax (‘synchronic’ analysis) in accordance to fossil-based evidence.

According to Donald (1991) and Deacon (1997), the establishment of syntax, with the “emergence of words as carriers of symbolic reference” (Knight et al., 2000, p. 8) as a foundation, relied on a prior high-speed phonetic machinery. Thus, the first step proto-humans made towards the establishment of language was through phonetics. When outlining the theory of evolution, Darwin himself attributed the origin of language “to the imitation and modification, aided by signs and gestures of various natural sounds, the voices of other animals, and man’s own instinctive cries (such as reaction to pain or surprise)” (1871, p. 56). Thus it was from a phonetic origin that language evolved. Some linguists agree with the theory that phonemes are the bare-bones of language. Hockett for example described phonemes as “the minimum meaningless but differentiating ingredients in a language; the smallest meaningful elements are what we call morphemes” (1959, p. 33). Further investigation deduced that increasing reliance on vocalisation was a focal point of human evolution, as it “freed the hands, allowing simultaneous foraging and enabled multiple partners to be ‘groomed at once’ whilst still communicating” (Dunbar, 1998, p. 101).

However, the early use of phonemes by proto-humans must be intertwined with the evolution of morphemes in language development – a distinct meaning cannot be communicated without a distinct phoneme or combination of phonemes.

### 3 Language Patterns Giving Evidence of the Dispersal of Human Communities

Recent research has made use of current phonetic diversity as a measure of linguistic development, based on assumptions about how phoneme use in a language evolves over time.

A controversial study by Atkinson (2011) takes tools used by biologists to trace genetic diversity and applies them to linguistic development, mixing the “serial founder effect” which was already observed in human biology with ‘phonetic diversity’ (Ramachandran et al., 2005). As populations migrate to remote locations, each episode of migration cuts down genetic diversity, and the community then settles for a time, the ‘serial founder’ then reproducing from a reduced genetic pool. It was therefore known that “human genetic and phenotypic diversity declines with distance from Africa” (Atkinson, 2011, p. 346), as dispersed communities had passed through “population bottlenecks” (Atkinson, 2011, p. 347). Applying this to phonemic diversity, Atkinson works on the assumption that “if phoneme distinctions are most likely lost in small founder population, then a succession of founder events during range expansion (dispersal) should progressively reduce phonemic diversity with increasing distance from the point of origin” (2011, p. 347). From a global sample of 504 languages, Atkinson demonstrates that African languages today have large phonemic inventories, while geographically distant languages “toward the far end of the human migration route out of Africa” (2011, p. 347), such as South American and Oceanian languages, have few: Hawaiian has only 13.

Statistical analysis linking phonemic diversity to dispersal patterns leads Atkinson to two conclusions. First, that the serial founder effect has parallel consequences on genetic and linguistic diversity as human communities spread beyond Africa, and secondly that the origin of language could be traced to Southwest Africa where phonemic diversity was greatest (since this is the community least affected by the serial founder effect).

Taking similar assumptions regarding the reasons for phoneme loss and even using Atkinson’s work as a source, Perreault and Mathew (2012) examined the potential that modern phonetics can reveal about modern human evolution and disparity. They observe that phonemic diversity evolves slowly, and can be used ‘as a clock’ to set a date for the origin of language, specifically in the presumed location of linguistic development in Southern Africa. Their analysis aims to establish how long it would take the oldest African languages would take to “accumulate the number of phonemes they possess today.” In the absence of anatomical or genetic data confirming when the physical capacity to speak evolved into actual language use, they claim this method can “circumvent problems that prevent current historical linguistics approaches from tackling the problem of dating the origin of language.” Like Atkinson, they observe that since human populations go through ‘bottlenecks’ when spreading out around the planet, phonemic diversity is lost. This founder effect means that Rotokas and Pirahã, in New Guinea and South America “both have 11 phonemes, while !Xun, a language spoken in Southern Africa has 141” (Perreault & Mathew, 2012). In turn, they argue that since phoneme loss takes a long time to recover, by comparing phonemic diversity against range of dispersal, it is possible to measure the age of languages. They create a timeline by comparing phonemic diversity in two environments settled at the same approximate time, one very isolated with almost no phonemic development, the Andaman Islands, and much more diverse envi-

ronment with a higher rate of phonemic evolution, Southeast Asia. They use this to calculate a rate of phonemic development, and then apply this measure to a stable linguistic group, in Africa, to produce an estimate that language originated there “in the Middle Stone Age (MSA) in Africa, sometime between 350–150 kya.” They argue that this method, comparing contemporary language phoneme diversity and tracking the relationship between human dispersal and phonemic change, corresponds with archeological evidence.

By their analysis, they conclude that language and phoneme diversity became “increasingly complex during the Middle Stone Age, between 350 kya-150 kya” rather than being a ‘spark’ that led to human colonization 50 kya, as the Southern African examples correspond to the described rate that settled languages maintain phonemes over time.

These two recent studies therefore suggest firstly that the origin of language was in Southern, not Eastern, Africa, and that it took place much earlier than other models suggested. The consequences for theories of human evolution and dispersal are great. It would confirm that language did co-evolve with the genetic evolution of *homo sapiens*, matching some of the evidence mentioned above about language in other species of the genus *homo*, and our pre-human ancestors. Additionally, it would mean that language originated only once, with its origins traced to Southern Africa. Apparently, this is confirmed by archeological evidence and the fossil record, which provide a physical link to support these mathematical conclusions.

A real strength of these investigations is how they correspond to prior genetic research and evidence on various levels. As well as relying on the overwhelmingly accepted consensus of an African origin for humanity, more local instances of genetic variety fit into the proposed theory. For instance, Oceanic populations native to Australia, the Trobriand Islands, and Papua New Guinea show enough genetic independence from their Asian forebears to demonstrate that they were the last ‘bottleneck’ migrants, in the same way the phonetic data does. On top of this, it explains later instances of genetic diversity, as “highest levels of phonemic diversity outside of Africa, are found in language families thought to be autochthonous to Southeast Asia” – another theory that Atkinson (2011, p. 347) claims fits into genetic evidence proposed by Palanichamy et al. (2004).

This analysis even harks back to the earlier work in language by early figures for establishing the modern theory of evolution, such as Darwin, in adapting the question of modern language to it (Knight et al., 2000). Most evidence in the references that Atkinson uses from genetic based research come to the same conclusion about our understanding of human dispersal and evolution. The method of working in parallel to genetic evidence to come up to linguistic conclusion (and indeed the nature of them) harks back to Chomsky: where language was one of the genetic ‘mutations’ that prompted the relentless spread of humans out of Africa, across the entire planet in a few thousand years.

## 4 Critical Review

The nature of the research into the origins of language is unavoidably based on theory and supposition, given the lack of directly available pieces of evidence for language going back any later than 10,000 years. Both studies acknowledge the uncertainty in this approach, and talk about the limitations in their methodologies. Perreault and Mathew (2012) say their research

is “built upon various assumptions that require further testing.” They mention assumptions that languages evolved at similar rates in different locations, and the bias that could be caused by the absence of tonal distinctions in the data they used. However, they claim their analysis “constitutes the first appraisal of when language evolved to be based directly on linguistic data” (2012).

Essentially, Atkinson’s conclusion is based on the mere observation that “human populations have lost phonemes through a drift-loss process during their expansion across the world” (Perreault & Mathew, 2012), as other rival hypotheses are yet to be rejected. The claim that phoneme development is more or less consistent with human development may seem to make sense in relation to fossil based evidence when examined in isolation with the above equations, yet it seems bold to suggest that elements such as later colonisation and conquest where certain languages would be held in higher social regard than others – even in Southern Africa. Bickerton goes on far enough to exclaim that “the biological evolution of humans has for all intents and purposes stopped”, compared to the rapidity of cultural change” (Bickerton, 2007, p. 511).

Another issue among Atkinson’s assumptions is that within the general rule that “the number of phonemes in a language is positively correlated with the size of its speaker population in such a way that small populations have fewer phonemes (2011, p. 346)” may correspond to the initial ‘spark’ of the bottleneck migration out of Africa. Yet, this rule fails to correspond to today’s languages and populations, even though Atkinson is using a set of data used only from modern languages. Applying such a population related rule extrapolated from modern language that cannot be applied to the current language-population accurately. However, the argument could be made that most modern population trends over the last 6,000 years are extremely rapid when compared to the timescale of early human migration and complex language was present from ancient times till now. Yet this brings up another problem of the distinction between ‘historical linguistics’ and ‘evolutionary linguistics’. This underestimation is best brought out by his confusion of the terms “phonemic diversity” and “phoneme inventory size”, wherein “diversity refers to variation within populations of individuals”, whereas Atkinson’s linguistic diversity refers to differences between languages (Atkinson, 2011, p. 657).

Any attempt to trace back linguistic development not only to pre-history but to the very emergence of *homo sapiens* as a species is going to be speculative if it is based only on current linguistic patterns. Each of the assumptions that form a basis for statistical analysis can be questioned, as the authors recognise. However, there is an undoubted link between these three sets of data – language, genetics, and dispersal patterns. Physical evidence concerning genetics and human dispersal – which is easier to date and easier to assign to geographical locations – does seem to corroborate the seemingly theoretical or statistical findings that these researchers derive from current language patterns.

Finally there are issues raised with the very linguistic data that both Atkinson’s study, and therefore Perreault and Matthew’s response, use. This data has been described as false and showing bias toward the heavy linkage between human and linguistic evolution that is described. Cysouw et al. (2012) argue that the unjustified weight in his data biases toward Africa’s “large phoneme inventories” (p. 11) as instead when the UPSID data are appropriately corrected for speaker community size and linguistic genera through a mixed-effects model (by

giving equal balance between vowel, tone and consonants), North America, as the area with the largest phoneme inventories, would be considered the ‘cradle of language’.

## References

- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from africa. *Science*, 332(6027), 346–349. <https://doi.org/10.1126/science.1199295>
- Bickerton, D. (2007). Language evolution: A brief guide for linguists. *Lingua*, 117(3), 510–526. <https://doi.org/10.1016/j.lingua.2005.02.006>
- Cysouw, M., Dediu, D., & Moran, S. (2012). Still no evidence for an ancient language expansion from africa. <http://www.sciencemag.org/content/335/6069/657.2.full>
- Darwin, C. (1871). The descent of man, and selection in relation to sex. *Nature*, 3(75), 442–444. <https://doi.org/10.1038/003442a0>
- Deacon, T. W. (1997). *The symbolic species – the co-evolution of language & the brain (paper): The co-evolution of language and the brain*.
- Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.
- Dunbar, R. (1998). *Grooming, gossip, and the evolution of language*. Harvard University Press.
- Hockett, C. F. (1959). Animal “languages” and human language. *Human Biology*, 31(1), 32–39. <http://www.jstor.org/stable/41449227>
- Knight, C., Studdert-Kennedy, M., & Hurford, J. (2000). *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511606441>
- Palanichamy, M. G., Sun, C., Agrawal, S., Bandelt, H.-J., Kong, Q.-P., Khan, F., Wang, C.-Y., Chaudhuri, T. K., Palla, V., & Zhang, Y.-P. (2004). Phylogeny of mitochondrial DNA macrohaplogroup n in india, based on complete sequencing: Implications for the peopling of south asia. *The American Journal of Human Genetics*, 75(6), 966–978. <https://doi.org/10.1086/425871>
- Pennisi, E. (2013). More genomes from denisova cave show mixing of early human groups. *Science*, 340(6134), 799–799. <https://doi.org/10.1126/science.340.6134.799>
- Perreault, C., & Mathew, S. (2012). Dating the origin of language using phonemic diversity (M. D. Petraglia, Ed.). *PLoS ONE*, 7(4), e35289. <https://doi.org/10.1371/journal.pone.0035289>
- Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa. *Proceedings of the National Academy of Sciences*, 102(44), 15942–15947. <https://doi.org/10.1073/pnas.0507611102>
- Schoenemann, P. T. (2009). Evolution of brain and language. *Language Learning*, 59, 162–186. <https://doi.org/10.1111/j.1467-9922.2009.00539.x>

- Tagliabue, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2008). Communicative signaling activates 'broca's' homolog in chimpanzees. *Current Biology*, 18(5), 343–348. <https://doi.org/10.1016/j.cub.2008.01.049>
- Valladas, H., Joron, J. L., Valladas, G., Arensburg, B., Bar-Yosef, O., Belfer-Cohen, A., Goldberg, P., Laville, H., Meignen, L., Rak, Y., Tchernov, E., Tillier, A. M., & Vandermeersch, B. (1987). Thermoluminescence dates for the neanderthal burial site at kebara in israel. *Nature*, 330(6144), 159–160. <https://doi.org/10.1038/330159a0>
- Wade, N. (2011). Phonetic clues hint language is africa-born. <https://www.nytimes.com/2011/04/15/science/15language.html>