This is an author produced version of a paper published in *Linguistics and Philosophical Investigations*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:


Abstract at publisher's web site:


Access to the published version may require subscription. Published with permission from: Addleton Academic Publishers (http://addletonacademicpublishers.com/authors-page/general/info/abstracts-linguistic-and-philosophical-investigations.html)
The precise timing of the emergence of language in human prehistory cannot be resolved. But the available evidence is sufficient to constrain it to some degree. This is a review and synthesis of the available evidence, leading to the conclusion that the time when speech in some form became important for our ancestors can be constrained to be not less than 400,000 years ago, thus excluding several popular theories involving a late transition to speech.

1. Introduction

That modern humans have language and speech, and that our remote ancestors did not, are two incontrovertible facts. But there is no consensus on when the transition from non-language to language took place, nor any consensus on whether it was a sudden jump or a gradual process. In this paper, I will explore to what extent the timing of the transition to language can be constrained by fossil, archeological, genetic, and other evidence. A fuller discussion of this and related issues can be found in Johansson (2005).

2. Upper limits

Very little can be said about upper limits on the age of language. Our closest relatives today do not have language in any reasonable sense, at least not in the wild. A reasonable inference from the lack of native ape language is that the last common ancestor of us and the other apes also lacked language. The alternative, that language evolved earlier and was subsequently lost in the chimpanzee lineage, appears implausible — why would something as useful as language be lost by a species heavily engaged in social communication?

The last common ancestor of humans and chimps almost certainly lived less than 10 million years ago (mya) (Benton & Donoghue, 2007), and conceivably as recent as 4 mya, though such a recent date is becoming difficult to reconcile with fossils. The current best estimate from molecular data is around 5 mya (Pääbo, 2003).

The oldest known fossils that with some confidence can be assigned to the human line are those in the genus *Ardipithecus*, from 4 - 6 mya, found in Ethiopia.
(Haile-Selassie et al., 2004; White et al., 2009). Other possible contenders for the earliest proto-human fossil are *Sahelanthropus tchadensis*, a skull found in Chad (Brunet et al., 2002), from about 6-7 mya with a puzzling mixture of features making it difficult to classify, and *Orrorin tugenensis* (Senut et al., 2001) just below 6 million years old, found in Kenya.

The discoverers of *Ardipithecus, Sahelanthropus* and *Orrorin* all consider their own fossil to be a human ancestor, and the others to be side branches (Cela-Conde & Ayala, 2003). I find the case for *Ardipithecus* somewhat more compelling, especially after the recent discovery of large parts of a skeleton (White et al., 2009), but the jury remains out. All these fossils have sufficient similarities with both humans and other apes that they are likely to be quite close to the branching point in the family tree. This would give an estimate from fossils of the most likely age of the last common ancestor somewhere in the vicinity of 6 mya, consistent within the uncertainties with the molecular estimate. Benton & Donoghue (2007) gives 6.5 mya as an estimate of the minimum time since the fossil split.

But beyond this common ancestor there is no way to place any more stringent upper limit on the time of language emergence. It may appear unlikely that e.g. australopithecines possessed language 2-3 mya, but there is no hard evidence that excludes it. Australopithecines appear to have lacked some speech adaptations, as discussed in the next section, having an ape-like hyoid bone (Alemseged et al., 2006) and thorax innervation (MacLarnon & Hewitt, 2004), but this does not exclude simpler forms of speech, much less non-speech language modalities. It has also been argued both that a minimum brain size is needed for language, and that the presence of language implies a human-like material culture. Both of these arguments are somewhat plausible, but it is by no means established that they exclude language in Lucy.

Thus, the best firm upper limit that can be placed on the time of language emergence remains a bit beyond 6 million years ago.

### 3. Lower limits

All modern human populations have language, obviously. Given that language has at least some biological substrate (if not necessarily an innate grammar) this implies that the most recent common ancestor of all modern humans had language, and had all the biological prerequisites for language.

The fossil record of anatomically modern *Homo sapiens* goes back to nearly 200,000 years ago in Africa (MacDougall et al., 2005; Marean, 2010). The molecular data likewise strongly support a common origin for all extant humans somewhere around 100,000 – 200,000 years ago (Ayala & Escalante, 1996; Wood, 1997; Bergström et al., 1998; Cavalli-Sforza & Feldman, 2003; Fagundes et al., 2007; Atkinson et al., 2008). The so-called ‘Mitochondrial Eve’ (Cann et al., 1987; Saville et al., 1998; Cavalli-Sforza, 1998), the putative common ancestress of all women, was at the forefront of this molecular wave, but she has since been
joined by a corresponding ‘Y-chromosome Adam’ (Fu et al., 1996; Dorit et al., 1995; Pääbo, 1995), somewhat more recent than Eve, as well as by data from non-sex-linked genes (Fischman, 1996), and from X-chromosomes (Disotell, 1999). The relation between population divergence times and genetic coalescence times is non-trivial (Hurford & Dediu, 2009), but it is hard to reconcile the genetic data with a common ancestor of all modern humans living much less than 100,000 years ago. It follows that the origins of the human language capacity cannot be more recent than 100,000 years ago.

This conclusion is consistent with archeological evidence of the spread of *Homo sapiens* out of Africa and the peopling of various continents, notably Australia 50,000 years ago or more (Hudjashov et al., 2007), and with evidence from morphology (Schillaci, 2008) and DNA (Macaulay, 2005) that indicates an ancient split between some Australasians and other Eurasian populations. In particular, Andaman islanders appear to have been genetically isolated for 65,000 years (Macaulay, 2005) and nevertheless have fully human language.

This 100,000-year lower limit already excludes theories that connect the origins of language with the apparent ‘cultural revolution’ in the archeological record around 40,000–50,000 years ago, discussed in the next section.

For at least some aspects of language, stronger time limits are possible. Our habitual use of speech is reflected in certain aspects of our anatomy, that can be studied in fossils. Speech adaptations can potentially be found in our speech organs, hearing organs, brain, and in the neural connections between these organs, as well as the genes controlling their development.

- **Speech organs.** The shape of the human vocal tract, notably the lowering of the larynx already in infancy, is very likely a speech adaptation, even though some other mammals, such as big cats, also possess a lowered larynx (Fitch, 2009). Speech would not be impossible with an ape vocal tract, but it would be less expressive, with fewer vowels available (Lieberman, 2008; de Boer & Fitch, 2010). Furthermore, the vocal tract in living mammals is quite flexible, and a resting position different from the human configuration does not preclude a dynamically lowered larynx, giving near-human vocal capabilities, during vocalizations (Fitch, 2009). The vocal tract itself is all soft tissue and does not fossilize, but its shape is connected with the shape of the surrounding bones, the skull base and the hyoid. Already *Homo erectus* had a near-modern skull base (Baba et al., 2003), but the significance of this is unclear (Fitch, 2000; Spoor, 2000), and other factors than vocal tract configuration, notably brain size and face size (Bastir et al., 2010), strongly affect skull base shape. Hyoid bones are very rare as fossils, as they are not attached to the rest of the skeleton, but one Neanderthal hyoid has been found (Arensburg et al., 1989), as well as two hyoids from *Homo heidelbergensis* (Martínez et al., 2008), all very similar to the hyoid of modern *Homo sapiens*, leading to the conclusion
that Neanderthals had a vocal tract adequate for speech (Houghton, 1993; Boë et al., 1999; Boë et al., 2007; Granat et al., 2007, but see also Lieberman (2007) and Fitch (2009)). The hyoid of *Australopithecus afarensis*, on the other hand, is more chimpanzee-like in its morphology (Alemseged et al., 2006), and the vocal tract that Granat et al. (2007) reconstruct for *Australopithecus* is basically apelike.

- **Hearing organs.** Some fine-tuning appears to have taken place during human evolution to optimize speech perception, notably our improved perception of sounds in the 2-4 kHz range. The sensitivity of ape ears has a minimum in this range, but human ears do not, mainly due to minor changes in the ear ossicles, the tiny bones that conduct sound from the eardrum to the inner ear. This difference is very likely an adaptation to speech perception, as key features of some speech sounds are in this region. The adaptation interpretation is strengthened by the discovery that a middle-ear structural gene has been the subject of strong natural selection in the human lineage (Olson & Varki, 2004). According to Martínez et al. (2004), these changes in the ossicles were present already in the 400,000-year-old fossils from Sima de los Huesos in Spain, well before the advent of modern *Homo sapiens*. These fossils are most likely Neanderthal ancestors, that Martínez et al. (2004) attribute to *Homo heidelbergensis*. In the Middle East, ear ossicles have been found both from Neanderthals and from early *Homo sapiens*, likewise with no meaningful differences from modern humans (Quam & Rak, 2008).

- **Brain.** Only the gross anatomy of the brain surface is visible as imprints on the inside of well-preserved fossil skulls. In principle, the emergence of e.g. Broca’s area could be pinpointed this way. But other apes have brain structures with the same gross anatomy as both Broca and Wernicke (Gannon et al., 1998; Cantalupo & Hopkins, 2001), so the imprints of such areas in the skulls of proto-humans tell us nothing useful about language. The rewiring of neural circuits that would tell us something useful, does not leave any fossil traces.

**Lateralization.** There is no clearcut increase in general lateralization of the brain in human evolution — ape brains are not symmetric (Balzeau & Gilissen, 2010) — and fossils are rarely undamaged and undistorted enough to be informative in this respect. But when tools become common, handedness can be inferred from asymmetries in the knapping process, the use-wear damage on tools, and also in tooth wear patterns (Uomini, 2009; Frayer et al., 2010), which may provide circumstantial evidence of lateralization, and possibly language (Steele & Uomini, 2009). Among apes there may be marginally significant handedness, but nothing like the strong
population-level dominance of right-handers that we find in all human populations. Evidence for a human handedness pattern is clear among Neanderthals and their predecessors in Europe, as far back as 500,000 years ago (Frayer et al., 2010), and some indications go back as far as 1 mya (Uomini, 2009). To what extent conclusions can be drawn from handedness to lateralization for linguistic purposes is, however, unclear.

- **Neural connections.** Where nerves pass through bone, a hole is left that can be seen in well-preserved fossils. Such nerve canals provide a rough estimate of the size of the nerve that passed through them. A thicker nerve means more neurons, and presumably improved sensitivity and control. The hypoglossal canal, leading to the tongue, has been invoked in this context (Kay et al., 1998), but broader comparative samples have shown that it is not useful as an indicator of speech (DeGusta et al., 1999; Jungers et al., 2003). A better case can be made for the nerves to the thorax, presumably for breathing control (Fitch, 2009). Both modern humans and Neanderthals have wide canals here, whereas *Homo ergaster* has the narrow canals typical of other apes (MacLarnon & Hewitt, 1999; MacLarnon & Hewitt, 2004), indicating that the canals expanded somewhere between 0.5 and 1.5 million years ago.

- **FOXP2.** When mutations in the gene FOXP2 were associated with specific language impairment (Lai et al., 2001), and it was shown that the gene had changed along the human lineage (Enard et al., 2002, but see also Diller & Cann, 2009), it was heralded as a “language gene”. But intensive research has revealed a more complex story, with FOXP2 controlling synaptic plasticity in the basal ganglia (Lieberman, 2009) rather than language per se, and playing a role in vocalizations and vocal learning in a wide variety of species, from bats (Li et al., 2007) to songbirds (Haesler et al., 2004). Nevertheless, the changes in FOXP2 in the human lineage quite likely are connected with some aspect of language, even if the connection is not as direct as early reports claimed. Relevant for the timing of the emergence of human language is that the derived human form of FOXP2 was shared with Neanderthals (Krause et al., 2007, but see also Benitez-Burraco et al. (2008) and Coop et al. (2008)), and that the selective sweep driving that form to fixation may have taken place more than a million years ago (Diller & Cann, 2009), well before the split between *Homo sapiens* and Neanderthals.

In conclusion, the fossil evidence indicates that at least some apparent speech adaptations were present in Neanderthals. No single one of these indications is compelling on its own, but their consilience strengthens the case for some form of Neanderthal speech. The presence of speech adaptations imply the presence of speech, or at least some form of complex vocalizations (though not vice versa — the absence of adaptations would not imply the absence of at least simple speech).
The apparent presence of speech in Neanderthals sets a lower limit for the age of speech at the time of the last common ancestor of us and the Neanderthals (unless one postulates major gene flow between the two lineages or, implausibly, the independent evolution of the same set of adaptations in both lineages). It has long been a controversial issue whether the Neanderthals actually were a separate lineage, or just a subspecies of Homo sapiens, and whether there was any significant admixture when modern humans replaced Neanderthals in Europe (Herrera et al., 2009). But a modest amount of admixture near the time of Neanderthal extinction would not materially affect the conclusions here, as the speech adaptations were in place in Neanderthals well before Homo sapiens invaded Europe and brought the two populations into proximity.

Early genetic evidence from fossil DNA in Neanderthals clearly supported their separateness (Hodgson & Disotell, 2008), and indicated that the last common ancestor lived at least 400,000 years ago (Krings et al., 1999; Höss, 2000; Beerli & Edwards, 2002; Knight, 2003; Caramelli et al., 2003; Endicott et al., 2010), though the limited statistics makes it impossible to exclude a modest level of admixture. The fossil evidence points in the same direction, with the earliest modern humans in Europe more resembling Africans than Neanderthals (Tyrrell & Chamberlain, 1998), though archaic features in some fossils may be interpreted as evidence of hybridization (Rougier et al., 2007; Trinkaus, 2007).

A draft sequence of the full Neanderthal genome was presented recently (Green et al., 2010), in which substantial similarities were found between the Neanderthal sequence and modern Eurasians. Green et al. (2010) interpret this as strong evidence of gene flow from Neanderthals into the common ancestor of modern Eurasians, but caution is in order as it is extremely difficult to exclude contamination with modern human DNA during excavation and processing of the fossils (Lalueza-Fox, 2009). Wall & Kim (2007) found evidence of both severe contamination and other problems in earlier work by Green et al. (2006). Firm conclusions should await replication.

There is no consensus on the taxonomy of the transitional fossils from around the time of our common ancestor with the Neanderthals. The names Homo heidelbergensis, rhodesiensis, antecessor, helmei and others are all in current use (Endicott et al., 2010). It is, however, quite well established that all of these have their roots in Homo erectus (sensu lato), so I will use the name erectus for our last common ancestor.

4. The revolution that wasn’t

The archeological record has frequently been invoked as support for the late, sudden appearance of language, due to the perception of a technological and creative revolution around 40,000–50,000 years ago (e.g. Binford (1989), Klein (1999), Li & Hombert (2002)).

Language use in itself is not archeologically visible, but other forms of sym-
bol use may be, and may be used as indicators that some level of semiotic abilities has been reached. Invoking ancient art, including pigments and personal ornaments, as indicators that the artists were capable of symbolic thought, or even as an indicator that language had evolved, is fairly common (Mellars, 1998; Henshilwood & Dubreuil, 2009). Some hunting techniques that require complex planning and mental time travel, such as snares and traps, may also be evidence of modern human cognition (Wadley, 2010). The precise connection between the decorative use of ochre and shell beads, and specific forms of language is, however, not worked out in detail, and not well supported (Botha, 2008).

The supposedly sudden appearance of advanced art and advanced tools in the caves of Europe about 40,000 years ago is often taken as evidence of a cognitive leap (Klein, 2008), indicating the origin of language. However, the appearance of a sudden dramatic ‘cultural revolution’ around 40,000 years ago, has turned out to be largely an illusion caused by the predominance of European sites in the documented archeological record, and possibly some Eurocentrism among archeologists (Henshilwood & Marean, 2003). *Homo sapiens* did indeed invade Europe rather suddenly about 40,000 years ago, bringing along an advanced toolkit — but that toolkit had been developed gradually in Africa over the course of more than 200,000 years (McBrearty & Brooks, 2000; Van Peer et al., 2003). Some aspects, such as blade technology, goes back as far 500,000 years (Johnson & McBrearty, 2010). Discoveries of works of abstract art (Henshilwood et al., 2002; Texier et al., 2010), pigment use (Barham, 2002; Henshilwood et al., 2009; Watts, 2009), and personal ornaments (Bouzouggar et al., 2007; Henshilwood & Dubreuil, 2009; d’Errico & Vanhaeren, 2009), all substantially older than 40,000 years, add further support to the long timescale of McBrearty & Brooks (2000). Also outside Africa some evidence of early pre-40k symbolic behavior has been found in the Middle East (Mayer et al., 2009).

The debate over the supposed revolution is reviewed by Bar-Yosef (2002) and Henshilwood & Marean (2003).

There was indeed a substantial and rapid increase in the frequency of modern human behavior in the Early Upper Paleolithic, but that increase may well be caused by demographic factors (Zilhão, 2007; Cartmill, 2010) rather than any cognitive changes. Similarly, ecological and energetic factors, rather than cognitive differences, may explain the sparsity of the Neanderthal artistic record, as evidenced by some undeniably modern human populations (e.g. Tasmanians) leaving archeological records resembling those created by Middle Paleolithic Neanderthals (Roebroeks & Verpoorte, 2009).

There is also possible evidence for simple art that actually predates the appearance of both Neanderthals and modern *Homo sapiens* (Bahn & Vertut, 1997; Bednarik, 2003), in the context of *Homo heidelbergensis* or possibly even *Homo erectus*.

Objects of the same type that are interpreted as evidence of symbolic behavior
when found in a *sapiens* context, have also been found associated with Neanderthals in Europe (Langley et al., 2008; d’Errico, 2008; Watts, 2009; Zilhão, 2007; Zilhão et al., 2010) as well as objects that can reasonably be interpreted as art (Appenzeller, 1998; d’Errico et al., 2003; Wynn & Coolidge, 2004). While these finds are simpler and less frequent than the ornaments and figurative art of later *Homo sapiens* (e.g. Bahn & Vertut (1997) and Conard (2003)), they nevertheless push back the origin of the biological capacities needed for symbolic behavior at least to the common ancestor of Neanderthals and us, some 400,000 years ago or more. And given that symbolic capacities are a prerequisite for language, and may be interpreted as indicative of the presence of language, this adds support to the case for Neanderthal language in some form, and for the time limit inferred from anatomy in the previous section.

5. Conclusions

Fossil evidence indicates that speech optimization of our vocal apparatus got started well before the emergence of *Homo sapiens*, almost certainly more than 400,000 years ago, probably in *Homo erectus*. As the speech optimization, with its accompanying costs, would not occur without strong selective pressure for complex vocalizations, presumably verbal communication, this implies that *Homo erectus* already possessed non-trivial language abilities.

There is no real evidence indicating just how complex language *erectus* had. It must have been complex enough to require fine-grained vocal distinctions, but this need not imply anything like modern grammar. They may have been at a holophrastic stage, or they may have had nearly full human language — it is difficult to imagine any way to tell. On one hand, *erectus* is the first hominin with a brain size approaching the modern human range — there are modern humans alive today with *erectus*-sized brains and excellent language skills — and they were also the first to spread out to many different habitats on different continents. But on the other hand their comparatively simple, static culture argues against their having modern human cognitive skills. In particular, it is quite clear that they lacked the cumulative cultural evolution that is so characteristic of modern humans. Given that they are different from modern humans in such fundamental ways, their having full modern human language appears unlikely.

Language need not have started in a spoken modality; sign language may have been the original language (e.g. Corballis (2002)), likely building on mimesis (Donald, 1997; Zlatev et al., 2005). This means language may be older than speech — but hardly younger. A lower age limit on speech remains a firm lower limit on the age of language at the *erectus* level, if not necessarily on full modern grammar.

Modern humans, after parting company with the Neanderthals perhaps close to half a million years ago, would have acquired the remaining features of modern language in parallel with acquiring modern human anatomy. Both aspects must
have been finished before modern humans started spreading over the world, perhaps 100,000 years ago. The last common ancestor of all humans today, probably living in Africa not so long before this exodus, is the likely speaker of Proto-World, the putative common ancestor of all the modern language families, and the earliest language which we may have any remote hope even in principle of ever reconstructing. But there is no reason to believe that this Proto-World was the first language spoken — as discussed above, our ancestors may have had language for a million years already. The details of those earlier proto-languages are likely to remain opaque for the foreseeable future.

References


Benitez-Burraco et al. (2008). Also sprach Neanderthalis or did she? *Biolinguistics* 2:2-3, 225-232


Boë & Maeda & Heim (1999). Neandertal man was not morphologically handicapped for speech. *Evolution of Communication* 3:49-77


d’Errico et al. (2003). Archaeological evidence for the emergence of language, symbolism, and music — an alternative multidisciplinary perspective. J World Prehistory 17:1-70


Fu et al. (1996). Estimating the age of the common ancestor of men from the ZFY intron. Science 272:1357-1365


Henshilwood et al. (2002). Emergence of modern human behavior: Middle Stone Age engravings from South Africa. *Science* 295:1278-1280

Henshilwood & d’Errico & Watts (2009). Engraved ochres from the Middle Stone Age levels at Blombos cave, South Africa. *J Hum Evo* 57:27-47


Hudjashov et al. (2007). Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis. *Proc Nat Acad Sci* 104:8726-8730


Krause et al. (2007). The derived FOXP2 variant was shared with Neandertals. *Current Biology* 17:1908-1912
Krings et al. (1999). DNA sequence of the mitochondrial hypervariable region II from the Neanderthal type specimen. Proc Nat Acad Sci 96:5581-5585


Li et al. (2007). Accelerated FoxP2 evolution in echolocating bats. PLOS One 2(9):PMID:17878935


Martínez et al. (2008). Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). J Hum Evo 54:118-124


Texier et al. (2010). A Howiesons Poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at Diepkloof Roch Shelter, South Africa. *Proc Nat Acad Sci* 107:6180-6185


Wadley, L. (2010). Were snares and traps used in the Middle Stone Age and does it matter? A review and a case study from Sibudu, South Africa. *J Hum Evo* 58:179-192


