TÜRK NÖROLOJİ DERGİSİ
TURKISH JOURNAL OF NEUROLOGY

Özgün Derleme Yazısı / *Original Review Article*

Türk Nöroloji Dergisi 2007; Cilt:13 Sayı:3 Sayfa:173-188

The Evolution of Language in Three Stages

Michael C. Corballis
University of Auckland, AUCKLAND

ABSTRACT
The Evolution of Language in Three Stages
Syntactic language is a uniquely human accomplishment, and must therefore have evolved since the split of the hominins from the other great apes some six million years ago. I argue that there were three main phases. The first came about through the emergence of bipedalism, a distinctively hominin trait, which enhanced the capacity for manual communication by freeing the hands and opening a frontal stance. The second began with the emergence of the larger-brained genus Homo from around 2 million years ago. The increase in brain size may have been driven by the necessity for enhanced social cooperation, and the emergence of a more effective system for communicating propositional information. Many of the properties of language, including the use of arbitrary symbols, and the emergence of tense and other markers of time and place, may have been driven by the increased understanding of time, and the advantages gained by recording and communicating episodic events. In short, language acquired syntax. The final stage, unique to Homo sapiens, was the emergence of autonomous speech as the primary mode of communication, replacing earlier dependence on manual gesture. This may help explain the dominance of our species over other hominin species, and indeed over the planet.

Keywords: evolution, language, gestural theory, FOXP2 gene

ÖZET
Hominin ve büyük maymunun altı milyon yıl önce birbirinden ayrılması ile gelişen, insanoğlunun en eşsiz başarısı, sözdizimsel dildir. Bu çalışmada dil gelişiminin üç ana aşaması ele alınmıştır. Birinci aşama, ellerin serbest kalmasını ve frontal duruş biçiminin oluşmasını sağlayarak el ile iletişimi geliştiren, tamamen insanımsı bir davranış olan bipedalizmin ortaya çıkmasıdır. İkinci aşama, yaklaşık iki milyon yıl önce büyük beyinli türlerden biri olan *Homo* türlerinin ortaya çıkışıdır. Beynin büyümesi ise, sosyal işbirliği gereksiniminin artması ve bilgiye dayalı iletişimin ortaya çıkması ile gelişmiş olmalıdır. Gelişigüzel sembollerin kullanılması, zamanların ve diğer yer-zaman belirteçlerinin ortaya çıkması gibi dilin birçok özelliği, zamanın giderek anlaşılmaya başlanması ve epizodik olayların kayıt edilmesi ve aktarılması ile kazanılmıştır. Kısacası, dil sözdizimini kazanmıştır. *Homo sapienlere* özgü son aşama ise, eski zamanlardaki el kol işareti ile anlaşmanın yerini alan, iletişimin ilk şekli olan kendine özgü anlatım biçiminin doğuşudur. Bu durum bizim, kendi türümüz, diğer insanımsı türler ve aslında gezegen üzerindeki hakimiyetimizi açıklamamıza yardımcı olabilir.

Anahtar kelimeler: evrim, dil, beden dili teorisi, FOXP2 geni

Yazışma Adresi/Address for Correspondence:
Michael C. Corballis
The University of Auckland - Private Bag 92019
Auckland Mail Centre AUCKLAND 1142 NEW ZEALAND
Tel: +64 9 373 7999
m.corballis@auckland.ac.nz

Kognitif IV Uluslararası Kognitif Nörobilim Sempozyumu'nda sunulmuştur.

INTRODUCTION

There is fairly general agreement that language is a uniquely human accomplishment. Although other species communicate in diverse ways, human language has properties that stand out as special. The most obvious of these is generativity -the ability to construct a potentially infinite variety of sentences, conveying an infinite variety of meanings. Animal communication is by contrast stereotyped and restricted to particular situations, and typically conveys emotional rather than propositional information. The generativity of language was noted by Descartes as one of the characteristics separating humans from other species, and has also been emphasized more recently by Chomsky, as in the following often-quoted passage:

"The unboundedness of human speech, as an expression of limitless thought, is an entirely different matter (from animal communication), because of the freedom from stimulus control; and the appropriateness to new situations... Modern studies of animal communication so far offer no counterevidence to the Cartesian assumption that human language is based on an entirely different principle. Each known animal communication system either consists of a fixed number of signals, each associated with a specific range of eliciting systems or internal states, or a fixed number of 'linguistic dimensions,' each associated with a non-linguistic dimension."[1]

More recently Chomsky, in a co-authored article, has acknowledged that some aspects of language may owe their origins to adaptations also present in other primates, but that the generative, recursive nature of language is distinctively human.*[2,3]

Despite general agreement on this point, there is controversy over how and when language evolved. Although captive chimpanzees, bonobos and other great apes have acquired some of the features of language, including the use of symbols to denote objects or actions, they have not displayed anything like recursive syntax, or indeed any degree of generativity beyond the occasional combining of symbols in pairs. To quote Pinker,[4] they simply don't "get it." This suggests that the common ancestor of humans and chimpanzee was almost certainly bereft of anything we might consider to be true language. Human language must therefore have evolved its distinctive characteristics over the past 6 million years. Some have claimed that this occurred in a single step, and recently -perhaps as recently as 170,000 years ago, coincident with the emergence of our own species. This is sometimes referred to as the "big bang" theory of language evolution. For example, Bickerton[5] asserted that "... true language, via the emergence of syntax, was a catastrophic event, occurring within the first few generations of *Homo sapiens sapiens* (p. 69)." Even more radically, Crow[6] has proposed that a genetic mutation gave rise to the speciation of *Homo sapiens,* along with such uniquely human attributes as language, cerebral asymmetry, theory of mind, and a vulnerability to psychosis.

Against this is a more conventional Darwinian view that language evolved in incremental fashion, through natural selection, as maintained by Pinker and Bloom[7] and elaborated by Jackendoff.[8] In this article, I propose that there were three broad phases to the emergence of modern language, spanning the past six million years.

Stage 1: Bipedal hominins

The journey toward articulate language began with the separation of the hominin line from that leading to modern chimpanzees and bonobos. The earliest fossil skull tentatively identified as a bipedal hominin is *Sahelanthropus tchadensis,* discovered in Chad,

* It has recently been claimed that starlings can parse recursive sequences, suggesting that phrase structure grammar may not be beyond the capacity of a nonhuman species. This claim is based on highly questionable evidence -see Corballis (in press).

and dated between 6 and 7 million years ago.[9] This date is probably very close to the time of the chimpanzee-hominin split, estimated at between 6.3 and 7.7 million years ago by a DNA-DNA hybridization technique.[10] Another early fossil, *Orrorin tugenensis,* is perhaps more securely identified as a hominin, and is dated from between 5.2 and 5.8 years ago.[11]

The feature that distinguished the hominins from the great apes was bipedalism. The idea that bipedalism may have contributed to the evolution of language was well anticipated by the 19th century linguist Wilhelm von Humboldt:

"And suited, finally, to vocalization is the upright posture of man, denied to animals; man is thereby summoned, as it were, to his feet. For speech does not aim at hollow extensions in the ground, but demands to pour freely from the lips towards the person addressed, to be accompanied by facial expression and demeanor *and by gestures of the hand,* and thereby to surround itself at once with everything that proclaims man human - *italics added.*"[12]

Although the upright stance might indeed have enhanced vocal communication, its more immediate and obvious effect would have been to free the hands. Nonhuman primates, no doubt including our immediate predecessors, also have a frontal stance, but are quadrupedal rather than bipedal. Our nearest primate relative, the chimpanzee, makes fairly extensive use of manual gestures, and these are arguably under more voluntary control than vocalizations, which in nonhuman primates serve as emotional signals rather than propositional utterances. Attempts to teach chimpanzees and other great apes to talk have proved fruitless,[13] but some progress toward human-like language has been accomplished using gestures, loosely based on American Sign Language,[14] or by having the animal point to symbols on a keyboard.[15] As noted above, the animals still don't "get" the concept of grammar, but appear able to use abstract symbols to denote objects and actions, and to combine them in pairs or triplets to make simple requests. Their vocabularies may number in the hundreds, but this is of course well short of the vocabulary attainable by humans, which numbers in the tens of thousands. Echoing Pinker, Bickerton[5] wrote that "The chimps' abilities at anything one would want to call grammar were next to nil," and has labeled this pre-grammatical level of linguistic performance "protolanguage."

Curiously, protolanguage has not been convincingly demonstrated among great apes in the wild, although closer study of natural gestures may well reveal greater sophistication than has hitherto been recorded. Nevertheless protolanguage may well form the platform from which true language evolved, as proposed by Jackendoff.[8] It is even possible that it was the discovery of protolanguage among the early hominins that drove bipedalism itself, freeing up a manual system ready-made for intentional action.

There are of course other possible explanations for the emergence of bipedalism, although none is completely satisfactory. One theory is that bipedalism evolved to allow more accurate throwing and clubbing.[16] Another controversial idea is that it was a product of an aquatic phase, in which the early hominins spent part of their existence wading in water, perhaps foraging for food in rivers, lakes, or the ocean side.[17-19] The idea that bipedalism might have been driven by selection for more effective communication is also controversial, especially since it implies that language itself evolved from manual gestures, rather than from vocal calls.

The gestural theory of language evolution

The idea that language may have evolved from manual gestures is nevertheless old, dating back at least to Condillac,[20] and was revived in a landmark article by Hewes.[21] It has been increasingly advocated over the past decade.[22-29] It is based on a number of considerations, and not merely on the manual communicative capacities of great apes. The late William C. Stokoe based his argument for the gestural theory primarily on the now well-established fact that the signed languages of the deaf have all the generativity and syntactic sophistication of spoken languages.[30,31] Four additional "facts," he claimed, show further that visible language must have preceded speech:

1. Sign languages still exist.
2. Spoken utterances often require visible signs in order to be fully understandable.
3. Only visible signs have natural links to concepts and syntactic structures.
4. All human infants use gesture to communicate before they master the language of their caretakers, whether that is a spoken language or a signed language.[29]

The theory has also received support from primate neurophysiology. Area F5 of the ventral premotor cortex of the monkey is the homologue of Broca's area, which in humans is involved in the production of vocal speech. In primates, however, area F5 has to do with manual action, not vocalization. Cells in F5 respond when the animal makes grasping movements with the hand or mouth, and a subset of those cells, dubbed "mirror neurons," also fire when the animal observes another individual making the same movements. It is now known that mirror neurons are part of a more general mirror system, involving areas in the superior temporal sulcus, and parietal lobe, as well as premotor cortex.[32] The direct mapping of perceived action onto the production of action seems to provide a platform for the evolution of language.[28] It is reminiscent of the so-called motor theory of speech perception, which postulates that we perceive speech with reference to its production, and not through acoustic analysis.[33] That is, the mapping system underling the perception of speech may be part of a more general system for the understanding of biological action.

There is so far no evidence that the mirror system is involved in the perception or production of vocalization in nonhuman primates, suggesting that the incorporation of vocalization into the system may have occurred late in hominin evolution -this is discussed further below. The vocalizations of nonhuman primates are probably largely automatic and emotional, controlled by the limbic system rather than the cortex.[34] Nevertheless there are cells in area F5 of the monkey that respond to the sounds of certain actions, such as the tearing of paper or the cracking of nuts.[35] That is, the mirror system may have been pre-adapted for the analysis of sound in terms of the actions that produced those sounds, but not for vocalization itself.

Stage 2: Large-brained *Homo*

The early bipedal hominins may well have communicated more effectively than their primate forebears, with the freeing of the hands and arms providing for a more extensive signaling capacity. There is little reason to suppose, however, that they had evolved a form of communication beyond protolanguage.

The emergence of a more grammatical, generative form of communication may have begun only with the emergence of the genus Homo, from around 2 million years ago. Although there is no direct evidence pertaining to language itself, a number of other changes from that time suggest the development of greater cognitive complexity. Stone tool industries have been dated from about 2.5 million years ago in Ethiopia,[36] and tentatively identified with *Homo rudolfensis*. However these tools, which belong to the Oldowan industry, were primitive, and some have suggested that *H. rudolfensis* and *H. habilis*, the hominin traditionally associated with the Oldowan, should really be considered australopithecines.[37] The true climb to humanity, and to language, probably began a little later, with the emergence of the larger-brained *Homo erectus* around 1.8 million years ago, and the somewhat more sophisticated Acheulian tool industry dating from around 1.5 million years ago.[38]

Bigger brains

But tool manufacture was not the only guide to the advance of cognition, and was probably not even the most telling one, since the Acheulian industry remained fairly static for over a million years, and even persisted into the culture of early *Homo sapiens* some 125,000 years ago.[39] A better indicator may be the increase in brain size. According to estimates based on fossil skulls, brain size increased from 457 cc in *Australopithecus africanus,* to 552 cc in H. habilis, to 854 cc in early *H. erectus* (also known as *H. ergaster*), to 1016 cc in later *H. erectus,* to 1552 in *H. neanderthalensis,* and back to 1355 cc in *H. sapiens.*[40] These values depend partly on body size, which probably explains why *H. neanderthalensis,* being slightly larger than modern humans, also had slightly larger brains, but the picture is clearly one of a progressive increase, first clearly evident in early *Homo.*

The increase in brain size corresponds at least approximately to the era known as the Pleistocene, usually dated from about 1.8 million years to about 10,000 years ago[41] -although it has been argued that it should be dated from as early as 2.58 million years ago,[42] which corresponds more closely to the emergence of the genus *Homo.* With the global shift to cooler climate after 2.5 million years ago, much of southern and eastern Africa probably became more open and sparsely wooded.[43] This left the hominins not only more exposed to attack from dangerous predators, such as saber-tooth cats, lions, and hyenas, but also obliged to compete with them as carnivores. The solution was not to compete on the same terms, but to establish what Tooby and DeVore[44] called the "cognitive niche," relying on social cooperation and intelligent planning for survival. As Pinker[45] put it, it became increasingly important to encode, and no doubt express, information as to "who did what to whom, when, where, and why." The problem is that the number of combinations of actions, actors, locations, time periods, implements, and so forth, that define episodes becomes very large, and a system of holistic calls to describe those episodes rapidly taxes the perceptual and memory systems. Syntax may then have emerged as a series of rules whereby episodic elements could be combined.

Language, memory, and time

Episodic memory may itself be a uniquely human endowment,[46,47] evolving as a subsystem of declarative memory to record specific events.[48] Declarative memory is contrasted with nondeclarative memory in that it is conscious, and can be "declared." The adaptiveness of episodic memory, in particular, may derive not so much from its role as a record of the past as from its potential in planning the future.[49,50] Indeed, people probably remember only a small fraction of actual past episodes,[51] and events are often remembered inaccurately, even to the point that people will claim with some certainty to have remembered events that did not in fact happen.[52,53] What remembered episodes do is provide a vocabulary of episodic components from which to plan events in the future, and perhaps to mentally rehearse them in anticipation.

Language may have evolved as a means of sharing information about episodes, whether actual or planned, thus vastly increasing our episodic vocabularies. Sometimes, people confuse their own experiences with those of others, and of course we humans seem to have an insatiable appetite for stories, whether told round the campfire, written in novels, or played in theatres or on television screens. Sometimes even memory and fiction are confused. During several of his political campaigns, former US President Ronald Reagan told the story of a wounded gunner whose plane was hit by anti-aircraft fire in World War II, and he could not eject from his seat. His commander comforted him, saying "Never mind, son, we'll ride it down together," and was posthumously awarded the Congressional Medal of Honor for heroism. It was later revealed that this was not a true event, but was a scene from the 1944 movie *A Wing and a Prayer.**[52]

* And of course if the story were true, neither Reagan nor anyone else could have known the details, since both the gunner and the commander were killed in the crash.

Through episodic memory, we have come to understand the very concept of time, which itself may be uniquely human.[49] Our remembered episodes give us a sense of the personal past, which is readily extrapolated to a personal future. The notion of time also extends beyond the bounds of our own lives, as we ponder the events of history and prehistory, or worry about the future of the planet. The understanding of time also leads to an understanding of death, and anxiety over what happens to us after we die may have been partly responsible for the emergence of religions that promise life after death.

Many of the properties of language, including syntax, emerge naturally from the requirements of conveying episodic memories. Past and future episodes are remote from the present in time, and typically in space as well, and also involve individuals and objects that are not physically present. In order to describe these episodes we therefore need ways of representing these absent components. Languages, whether signed or spoken, accomplish this through the use of symbols. Of course nonhuman animals have also evolved ways of communicating symbolically. For example, vervet monkeys give different warning cries to distinguish between a number of different threats, such as snakes, hawks, eagles, or leopards. When a monkey makes one of these cries, the troop acts appropriately, clambering up trees in response to a leopard call or running into the bushes in response to an eagle call.[54] These cries bear no obvious relation to the sounds emitted by the predators they stand for, and are in that sense symbolic. Similarly, chimpanzees in the wild emit a pant-hoot call on the discovery of food,[55] and the bonobo Kanzi has learned hundreds of abstract symbols representing objects and actions.[15] These symbols refer to events in the present, whereas the symbols used by humans have extended properties of reference, involving continuity across time and space.[56]

One of the features of syntax is tense, which has to do with conveying information about time itself. Thus we can refer to events that happened in the past, or to anticipated future events. Indeed, syntactic structures go much further and allow us to specify what will have happened, what might have happened, what was happening, what should happen, what was about to happen, and so forth. We have also developed semantic systems to locate events precisely in time, by specifying the century, decade, year, month, day, and even time of day. Displacement in time also usually means displacement in space; the events of yesterday typically happened in some place other than the present location, so we need conventions to refer to space. Indeed, it has been argued that the structure of language itself derives from our concepts of space,[57,58] and that time can also be represented spatially—the so-called "fourth dimension." In American Sign Language the time line runs from behind the body, representing the past, to the front of the body, representing the future.[30] Most prepositions in English, such as *at, about, around, between, among, along, across, against, from, to,* and *through*, are fundamentally spatial but can be used to refer to time as well, and only a few, such as *since* or *until*, apply specifically to the time dimension. Spatial prepositions are also transported to logical expressions, such as *A follows from B*, or *The argument against A is B*. The intimate link between grammar and space may derive from human development, and Mandler[59] has argued that language is structured spatially because preverbal infants have already constructed spatial representations, and syntax is then built onto these representations. The developmental links between language and spatial concepts is explored more fully in the edited volume by Bloom et al.[60]

In neuroanatomical terms, the key to the evolution of both mental time travel and language may lie in the mesial temporal lobe, and more particularly the hippocampus. O'Keefe and Nadel[61] showed that the infrahuman hippocampus comprises a "cognitive map," in which particular locations are stored for future use. O'Keefe suggests that this model must be extended to account for evidence on the role of the hippocampus in humans, in two ways. First, time itself must be incorporated into the model to allow for the storage of spatiotemporal information, thus providing for episodic memory. It is likely that the frontal lobes are also critically involved in adding the time dimension.[62] Second, the model must be extended to account for lateralization of function, with much of the left hemisphere taken over by language functions. In humans, damage to the right mesial temporal lobe results in amnesia for episodic visuospatial material, while left-sided damage results in amnesia for linguistic material.[63] Thus the right hippocampus receives inputs from analysis of the physical world, while the left hippocampus receives inputs from language centers. Again, this account emphasizes the priority of spatial structures in the representation of both time and language.

The complementary nature of language and space extends also to the neocortex. Language is represented in many areas of the brain, predominantly but not exclusively in the left hemisphere.[64] In some areas, notably in parietal and temporal lobes, the corresponding areas on the right side have to do with spatial perception,[65] suggesting that language usurped left-sided areas that had evolved earlier to subserve spatial functions. It should be noted, though, that cerebral asymmetry *per se* is not unique to our species, but has been widely documented in many if not most other species, including reptiles, birds, fish, mammals, and primates.[66] The seeds for lateralization were undoubtedly sown early in vertebrate evolution, but lateralization appears to be especially pronounced expression in the human brain, especially with respect to handedness, manual action, and visuospatial attention.

The development of episodic memory, mental time travel, and language probably began with the emergence of the genus *Homo*. There is also evidence that the spatial world of our forebears expanded during the Pleistocene. *Homo erectus* marked the progression from a relatively primitive form of bipedalism, retaining a degree of adaptation to an arboreal habitat, to the full striding gait characteristic of modern humans. It has been suggested that the newer form of bipedalism evident from *erectus* on was also an adaptation to efficient endurance running, and resulted in marked changes to skeletal structure.[67] From about 1.6 million years ago, some members of this species strode out of Africa and into Asia, and *erectus* fossils in Java have been dated to as recently as 30,000 years ago.[68]

In summary, the genus *Homo* was characterized by a greatly extended habitat, a brain structured by enhanced spatial and temporal understanding, and an unprecedented ability to communicate that understanding to others.

Stage 3: Articulate humans

Although syntactic language may have evolved during the Pleistocene, it was probably primarily manual rather than vocal, at least until the emergence of *Homo sapiens*. It may have resembled the signed languages invented by present-day deaf communities, although it is likely that vocal elements were increasingly introduced. Fossil evidence suggests that articulate speech emerged late in the evolution of the genus *Homo*, which gives further grounds for supposing that it may have been preceded by a gestural system. Indeed, I shall argue that autonomous speech did not emerge until the arrival of our own species, *Homo sapiens*, some 170,000 years ago, and perhaps even later, and may have been the distinctive characteristic that led to human domination on the planet, and the eventual extinction of all other hominin species.

The idea that language could have switched from a manual to a vocal form is nevertheless one of the main difficulties associated with the gestural theory.

In his 2005 book *The Talking Ape*, the linguist Robbins Burling[69] writes as follows:

... the gestural theory has one nearly fatal flaw. Its sticking point has always been the switch that would have been needed to move from a visual language to an audible one (p. 123).

The point is a fair one. On the surface, at least, it does indeed seem unlikely that language could have shifted from a system involving visual perception of hand and arm movements to one involving auditory perception of movements of the tongue, lips, and larynx. A deeper analysis, though, suggests continuity between hand and mouth, implying that the transition may have been relatively smooth.

The transition from hand to mouth

The transitional problem is eased by increasing recognition that speech itself is a gestural system, rather than an acoustic one. Earlier, I referred to the motor theory of speech perception, and the implication that speech is perceived, not through regular acoustic analysis, but through understanding of the actions producing the speech sounds. This has led to what is known as *articulatory phonology*,[70] in which speech is understood as comprised of articulatory gestures rather than of phonemes. Six articulatory organs -namely, the lips, the velum, the larynx, and the blade, body, and root of the tongue- produce these gestures. Each is controlled separately, so that individual speech units are comprised of different combinations of movements. The distribution of action over these articulators means that the elements overlap in time, which makes possible the high rates of production and perception.[71]

Given that speech is a gestural system, the transition from hand to mouth could well have been gradual, and in any case language has probably always been a combination of manual, facial, and vocal elements. Indeed it remains so today, since people characteristically gesture manually and facially as they speak, and these visible gestures can convey critical information.[72-74] Manual movements can also influence the actual production of speech. Relative to grasping a small object, grasping a large object[75] and bringing it to the mouth[76] induces selective increases in parameters of lip kinematics and voice spectra of syllables pronounced simultaneously with action execution. Even observing another individual grasping or bringing to the mouth larger objects affects the lip kinematics and the voice spectra of syllables simultaneously pronounced by the viewer.[77] These findings suggest a mechanism of double command to hand and mouth, which may have helped mediate transfer of the language medium from a manual to a vocal system.[78]

The evolutionary sequence of events may be paralleled by those in the development of language in children, in which gestures are intimately tied to vocalizations.[64] For example, canonical babbling in children aged from 6-8 months is accompanied by rhythmic hand movements. Manual gestures predate early development of speech in children, and predict later success even up to the two-word level. Word comprehension in children between 8 and 10 months and word productions between 11-13 months are accompanied by gestures of pointing and showing and gestures indicating recognition, respectively. Even in adults, it is well known that manual gestures accompany speech, to form a single integrated system.[74]

It is also likely that movements of the face were transitional between manual and vocal communication. Visible movements of the face can also influence the perception of speech, as in the McGurk effect, in which dubbing sounds onto a mouth that is saying something different alters what the hearer actually hears.[79] Watching speech movements, and even stills of a mouth making a speech sound, activate the mirror system, including Broca's area.[80] Although we can communicate without having to see the person we are talking to, as on radio or cell-phone, speech in the natural world is rendered more eloquent and meaningful with the addition of bodily movements.

What has changed, then, is the relative contribution of visual and vocal components. The early hominins, as we have seen, were naturally preadapted to communicate using the hands and arms, and any vocal contribution is likely to have been restricted to emotional grunts, perhaps somewhat resembling those of modern tennis players. Unlike vocalization, manual movements were under voluntary control, and free movement of the hands in space provided a natural theater for communicating information about events in the spatiotemporal world. Some degree of communication could be accomplished through mime, but this would become increasingly unwieldy with the increase in vocabulary size.

It is often claimed that signed languages are fundamentally iconic, in which objects or actions are sculpted by the hands as pictorial representations. Although some signs do have some degree of iconicity, most signs have lost their iconic component and become arbitrary. This requires a process of *conventionalization*, whereby the meanings of symbols become known to the linguistic community.[81] Conventionalization is as much a requirement of signed languages as of spoken ones, and different signed languages can be as "foreign" to one another as are different spoken languages. The use of arbitrary rather than iconic or analog representations may not be a fundamental property of language so much as a matter of expediency. Even in the visual domain, it is time-consuming to construct spatial representations of all objects and actions that we know, and in the speech domain there is little opportunity for analog representation, except perhaps in the case of onomatopoeia (as in words like buzz or swish). The switch to arbitrary representation also allows the development of symbols to refer to abstract ideas, like honesty or altruism.

Through evolutionary time, though, there would have been pressure to shift from the hands to the face, and ultimately to voicing. As *Homo* began to develop more sophisticated manufacturing techniques, there would have been increased competition between use of the hands for communication, and the use and manufacture of tools. Communication may therefore have increasingly incorporated facial movements, which are also under some degree of voluntary control, thereby freeing the hands for manufacture, as well as for carrying things. It is increasingly recognized that facial movements play an important role in present-day signed languages, and are critical to syntax. For example, in American Sign Language, interrogation is signaled by raising the eyebrows, and negation by shaking the head.[30]

Incorporation of mouth movements, in particular, would be facilitated by pre-existing neural connections having to do with the role of the hands in bringing food to the mouth.[82] Thus there may have been increasing pressure for the tongue, lips, and vocal tract to assume more of the communicative burden. Since the tongue, velum, and larynx are for the most part invisible, there may have been pressure to add sound, so that gestures of the mouth were rendered accessible. Adding voicing to the gestures also provides for the distinction between voiced and unvoiced sounds, adding to the possible repertoire. In this view, speech itself may be considered to be facial gesture, half swallowed. We are able to recover speech gestures somewhat in the way that monkeys, through the activity of the mirror system can recover, the actions leading to the sound of paper tearing, or nuts being cracked open.

Fossil evidence

Fossil evidence suggests that the anatomical changes necessary for articulate speech occurred gradually in evolution, and were probably not complete until the emergence of our own species, *Homo sapiens*. If this is so, then articulate speech was probably predated by the evolution of syntactic language itself. One piece of evidence has to do with the hypoglossal canal at the base of the tongue. The hypoglossal nerve, which passes through this canal and innervates the tongue, is much larger in humans than in great apes, probably because of the important role of the tongue in speech. Fossil evidence suggests that the size of the hypoglossal canal in early australopithecines, and perhaps in *Homo habilis*, was within the range of that in modern great apes, while that in Neanderthal and early *H. sapiens* was well within the modern human range,[83] although this has been disputed.[84] A further clue comes from the finding that the thoracic region of the spinal cord is relatively larger in humans than in nonhuman primates, probably because breathing during speech involves extra muscles of the thorax and abdomen. Fossil evidence indicates that this enlargement was not present in the early hominins or even in *Homo ergaster*, dating from about 1.6 million years ago, but was present in several Neanderthal fossils.[85,86]

The production of articulate speech in humans depends on the lowering of the larynx. According to P. Lieberman[87,88] this adaptation was incomplete even in the Neanderthals of 30,000 years ago, and their resultant poor articulation would have been sufficient to keep them separate from *H. sapiens*, leading to their eventual extinction. This work remains controversial,[89] but there is other evidence that the cranial structure underwent changes subsequent to the split between anatomically modern and earlier "archaic" *Homo*, such as the Neanderthals, *Homo heidelbergensis*, and *Homo rhodesiensis*. One such change is the shortening of the sphenoid, the central bone of the cranial base from which the face grows forward, resulting in a flattened face.[87] D. E. Lieberman speculates that this is an adaptation for speech, contributing to the unique proportions of the human vocal tract, in which the horizontal and vertical components are roughly equal in length. This configuration, he argues, improves the ability to produce acoustically distinct speech sounds, such as the vowel [i].[90] It is not seen in Neanderthal skeletal structure,[91] suggesting that it emerged in our own species within the past 500,000 years.

Another adaptation unique to *H. sapiens* is neurocranial globularity, defined as the roundness of the cranial vault in the sagittal, coronal, and transverse planes, which is likely to have increased the relative size of the temporal and/or frontal lobes relative to other parts of the brain.[92] These changes may reflect more refined control of articulation and also, perhaps, more accurate perceptual discrimination of articulated sounds.

Perhaps the most critical evidence that the switch to articulate speech emerged only in our own species comes, not from fossils, but from genetics.

The FOXP2 gene

About half of the members of three generations of an extended family in England, known as the KE family, are affected by a disorder of speech and language. The disorder is evident from the affected child's first attempts to speak and persists into adulthood.[93] The disorder is now known to be due to a point mutation on the FOXP2 gene (forkhead box P2) on chromosome 7.[94,95] For normal speech to be acquired, two functional copies of this gene seem to be necessary.

The nature of the deficit in the affected members of the KE family, and therefore the role of the FOXP2 gene, have been debated. Some have argued that FOXP2 gene is involved in the development of morphosyntax,[96] and it has even been identified more broadly as the "grammar gene"[4] -although Pinker[45] has since recognized that other genes probably played a role in the evolution of grammar. Subsequent investigation suggests, however, that the core deficit in affected members of the KE family is one of articulation, with grammatical impairment a secondary outcome.[97] It may therefore play a role in the incorporation of vocalization into the mirror system. This is supported by a study in which fMRI was used to record brain activity in both affected and unaffected members of the KE family while they covertly generated verbs in response to nouns.[98] Whereas unaffected members showed the expected activity concentrated in Broca's area in the left hemisphere, affected members showed relative underactivation in both Broca's area and its right-hemisphere homologue, as well as in other cortical language areas. They also showed overactivation bilaterally in regions not associated with language. However, there was bilateral activation in the posterior superior temporal gyrus; the left side of this area overlaps Wernicke's area, important in the comprehension of language. This suggests that affected members may have generated words in terms of their sounds, rather than in terms of articulatory patterns. Their deficits were not attributable to any difficulty with verb generation itself, since affected and unaffected members did not differ in their ability to generate verbs covertly. Another study based on structural MRI showed morphological abnormalities in the same areas in the affected members of the family.[99]

The FOXP2 gene is highly conserved in mammals, and in humans differs in only three places from that in the mouse. Nevertheless, two of the three changes occurred on the human lineage after the split from the common ancestor with the chimpanzee and bonobo. A recent estimate of the date of the more recent of these mutations suggests that it occurred "since the onset of human population growth, some 10,000 to 100,000 years ago."[100] Enard et al. further suggest that the date "is compatible with a model in which the expansion of modern humans was driven by the appearance of a more-proficient spoken language" (p. 871). What made language more proficient, then, may have been the final accomplishment of autonomous speech.

It is unlikely, though, that the FOXP2 mutation was the only event in the transition to speech, which undoubtedly went through several steps and involved other genes.[101] Moreover, the FOXP2 gene is expressed in the embryonic development of structures other than the brain, including the gut, heart, and lung.[102] It may have even played a role in the modification of breath control for speech.[85] A mutation of the FOXP2 gene may nevertheless have been the most recent event in the incorporation of vocalization into the mirror system, and thus the refinement of vocal control to the point that it could carry the primary burden of language.

The idea that the critical mutation of the FOXP2 gene occurred less than 100,000 years ago is indirectly supported by recent evidence from African click languages. Two of the many groups that make extensive use of click sounds are the Hadzabe and San, who are separated geographically by some 2000 kilometers, and genetic evidence suggests that the most recent common ancestor of these groups goes back to the root of present-day mitochondrial DNA lineages, perhaps as early as 100,000 years ago.[103] This could mean that clicks were a prevocal way of adding sound to facial gestures, prior to the FOXP2 mutation. Evidence from mitochondrial DNA suggests that modern humans outside of Africa date from groups who migrated from Africa and these groups may have already developed autonomous speech, leaving behind African speakers who retained click sounds. The only known non-African click language is Damin, an extinct Australian aboriginal language. *Homo sapiens* may have arrived in Australia as early as 60,000 years ago, not long after the migrations out of Africa.

In a recent review, Mellars[104] suggests that modern humans may have reached Malaysia and the Andaman Islands as early as 60,000 to 65,000 years ago, with migration to Europe and the Near East occurring from western or southern Asia, rather than from Africa as previously thought. This is not inconsistent with an estimate by Oppenheimer[105] that the eastward migration out of Africa took place around 83,000 years ago. Another recent study suggests that there was back-migration from to Africa at around 40,000 to 45,000 years ago, following dispersal first to Asia and then to the Mediterranean.[106] These dates are consistent with the view that autononomous speech emerged prior to the migration of anatomically modern humans out of Africa, and was indirectly responsible for the subsequent dominance of our species on the planet.

Why speech?

According to the account presented here, the transition from manual to vocal language was not abrupt. This raises the question, though, of why the transition took place at all. The signed languages of the deaf clearly show that manual languages can be as sophisticated as vocal ones, suggesting that the switch was not driven by linguistic considerations. There were also costs, since the transition to speech involved the lowering of the larynx, which greatly increased the risk of choking to death. Clearly, the evolutionary pressure toward speech must have been strong.

I have already suggested that the transition from hand to mouth may have been driven in part by the development of manufacture, and competition for the hands. Indeed vocal language allows people to speak and use tools at the same time, leading perhaps to pedagogy.[25] But this was probably not the only factor contributing to the pressure for change. For one thing, speech is much less energy-consuming than manual gesture. Anecdotal evidence from courses in sign language suggests that the instructors require regular massages in order to meet the sheer physical demands of sign-language expression. In contrast, the physiological costs of speech are so low as to be nearly unmeasurable.[107] In terms of expenditure of energy, speech adds little to the cost of breathing, which we must do anyway to sustain life.

Speech is also less attentionally demanding than signed language; one can attend to speech with one's eyes shut, or when watching something else. Speech also allows communication over longer distances, as well as communication at night or when the speaker is not visible to the listener. The San, a modern hunter-gatherer society, are known to talk late at night, sometimes all through the night, to resolve conflict and share knowledge.[108] Boutla et al.[109] have shown that the span of short-term memory is shorter for American Sign Language than for speech, suggesting that voicing may have permitted longer and more complex sentences to be transmitted -although the authors claim that the shorter memory span has no impact on the linguistic skill of signers.

The advantages of speech over manual gesture may even help explain why humans eventually predominated over other large-brained hominins, including the Neanderthals, who died out some 30,000 years ago. By the same token, the freeing of the hands and the emergence of pedagogy may help explain the so-called "human revolution",[110] manifest in the dramatic appearance of more sophisticated tools, bodily ornamentation, art, and perhaps music, dating from some 40,000 years ago in Europe, and probably earlier in Africa.[105,111] Changes in the mode of communication can have a dramatic influence on human culture, as illustrated by the invention of writing, and more recently by email and the Internet. These changes were relatively sudden, and cultural rather than biological. The change from manual to vocal communication, in contrast, would have been slow, driven by natural selection and involving biological adaptations, but it may have had no less an impact on human culture - and therefore, perhaps, on human fitness.[112]

Summary and conclusions

Among present-day species, fully grammatical language appears to be a uniquely human accomplishment. Other animals are capable of understanding symbolic representations, and perhaps even of segmenting speech, at least to the point of isolating words. Besides the bonobo,[15] this may include the Africa gray parrot[113] and the domestic dog.[114] But there is no evidence that nonhuman animals can decode or generate grammar, and so create and understand a potentially infinite variety of sentences. At best, they are at the level of the two-year-old human child, with a level of communication lacking the generative, recursive property of fully developed language. At best, the linguistic capacities of other species, like those of the two-year-old child, constitute protolanguage.

The emergence of language from protolanguage may have occurred late in hominin evolution, though not so late as to represent an evolutionary "big bang." The steps toward grammar may have begun some 2 million years ago, with the emergence of larger brained hominins, and continued over the next 1.5 million years, or thereabouts. The understanding of how this occurred has been described as the hardest problem in science.[115] I have suggested in this article that it was driven, at least in part, by the emergence of episodic memory and the discovery of time, and the adaptive advantages to be gained by sharing our experiences and plans. Fully syntactic language may have been achieved by some 500,000 years ago, when brain size leveled off to about the modern range.

The emergence of speech as an autonomous system probably occurred even later, and may not have been complete even in the Neanderthals, who survived until some 30,000 years ago. Indeed may not have been complete in our own species until less than 100,000 years ago, with the mutation of the FOXP2 gene allowing vocalization to be incorporated into the mirror system. Evidence from mtDNA suggests that modern humans migrated out of Africa some 83,000 years ago,[105] eventually replacing all other hominins, including the Neanderthals in Europe, *H. erectus* in Asia, and even groups of *H. sapiens* who had migrated earlier. What was it that led to the dominance of these late migrants? I have suggested that it may have been the consequences of the emergence of fully articulate speech, resulting in improved technology, perhaps including more lethal weaponry, and a more coherent culture.[112]

One might have thought that an understanding of how language evolved would have been beyond the reach of science. That, presumably, was the view in 1866, when the Linguistic Society of Paris banned all discussion of the topic. Nevertheless the past decade, in particular, has produced an extraordinarily rich accumulation of evidence from multiple sources, all of which appear to be converging on common themes, if not yet on an agreed scenario. In 1866, very little was known about the transitions from ape to human, but modern archaeology has given us a remarkably detailed account of what our hominin forebears must have been like. From sceptical talk of a "missing link" we now have evidence of over 20 hominin species separating us from our common ancestry with the chimpanzee and bonobo.[37] Detailed inspection of hominin fossils has provided evidence of brain size, and growth characteristics, and modern biochemistry has elucidated the timing of critical events, such as the ape-hominin split, and the late migration out of Africa.

We also now understand much better what language is actually like, how it differs from other forms of communication, and how it develops. It has only recently become clear that the signed languages of the deaf are true grammatical languages, and not impoverished signaling systems. With the advance of brain imaging, the neurophysiology of language is increasingly understood, and work on the so-called mirror system has led to important insights as to how language might be better understood as part of a more general system for understanding biological motion, instead of a rather abstract coding system beyond any affinity with our animal heritage. It can be safely anticipated that further discoveries and insights will contribute to the understanding of how language evolved, and no doubt modify many of the claims made in this article.

REFERENCES

1. Chomsky, N. (1966). Cartesian linguistics: A chapter in the history of rationalist thought. New York: Harper & Row.
2. Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? Science, 298, 1569-1579.
3. Corballis, M. C. (in press). Language, recursion, and starlings. Cognitive Science.
4. Pinker, S. (1994). The language instinct. New York: Morrow.
5. Bickerton, D. (1995). Language and human behavior. Seattle, WA: University of Washington Press.
6. Crow, T. J. (2002). Sexual selection, timing, and an X-Y homologous gene: Did Homo sapiens speciate on the Y chromosome? In T. J. Crow (Ed.), The speciation of modern Homo sapiens (pp. 197-216). Oxford, UK: Oxford University Press.
7. Pinker, S. & Bloom, P. (1990). Natural language and natural selection. Behavioral and Brain Sciences, 13, 707-784.
8. Jackendoff, R. (2002). Foundations of language: Brain, meaning, grammar, evolution. Oxford, UK: Oxford University Press.
9. Brunet, M., Guy, F., Pilbeam, D., Mackaye, H. T., Likius, A., Ahounta, D., Beauvilain, A. et al. (2002). A new hominin from the Upper Miocene of Chad, Central Africa. Nature, 418, 145-151.
10. Sibley, C. G., & Ahlquist, J. E. (1984). The phylogeny of hominoid primates, as indicated by DNA-DNA hybridisation. Journal of Molecular Evolution, 20, 2-15.
11. Galik, K., Senut, B., Pickford, M., Gommery, D., Treil, J., Kuperavage, A. J., & Eckhardt, R. B. (2004). External and internal morphology of the BAR 1002'00 Orrorin tugenensis femur. Science, 305, 1450-1453.
12. Humboldt, W. von. (1836). On language: The diversity of human language-structure and its influence on the mental development of mankind (trans. P. Heath). Cambridge: Cambridge University Press.
13. Hayes, C. (1952). The ape in our house. London: Gollancz.
14. Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. Science, 165, 664-672.
15. Savage-Rumbaugh, S., Shanker, S.G., & Taylor, T. J. (1998). Apes, language, and the human mind. New York: Oxford University Press.
16. Marzke, M. W. (1996). Evolution of the hand and bipedality. In A. Lock & C. R. Peters (Eds.), Handbook of symbolic evolution (pp. 126-154). Oxford: Oxford University Press.
17. Hardy, A. (1960). Was man more aquatic in the past? New Scientist, 7, 642-645.
18. Morgan, E. (1982). The aquatic ape. London: Souvenir Press.
19. Tobias, P. V. (1998). Water and human evolution. Out There, 35, 38-44.
20. Condillac, E.B. de (1971). An essay on the origin of human knowledge. T. Nugent (Tr.), Gainesville, FL: Scholars Facsimiles and Reprints. (Originally published 1746).
21. Hewes, G.W. (1973). Primate communication and the gestural origins of language. Current Anthropology, 14, 5-24.
22. Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. Behavioral & Brain Sciences, 28, 105-168.
23. Armstrong, D.F. (1999). Original signs: gesture, sign, and the source of language. Washington, DC: Gallaudet University Press.
24. Armstrong, D.F., Stokoe, W.C., & Wilcox, S.E. (1995). Gesture and the nature of language. Cambridge: Cambridge University Press.
25. Corballis, M. C. (2002). From hand to mouth: the origins of language. Princeton, NJ: Princeton University Press.
26. Givòn, T. (1995). Functionalism and grammar. Philadelphia, PA: John Benjamins.
27. Kelly, S. D., Iverson, J. M., Terranova, J., Niego, J., Hopkins, M., & Goldsmith, L. (2002). Speech and gesture on three time frames. Developmental Neuropsychology, 22, 323-349.
28. Rizzolatti, G. & Arbib, M. A. (1998). Language within our grasp. Trends in Cognitive Sciences, 21, 188-194.
29. Stokoe, W. C. (2001). Language in hand: Why sign came before speech. Washington, DC: Gallaudet University Press.
30. Emmorey, K. (2002). Language, cognition, and brain: Insights from sign language research. Hillsdale, NJ: Erlbaum.
31. Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). The syntax of American Sign Language. Cambridge, MA: The MIT Press.
32. Rizzolatti, G., Fogassi, L., & Gallese V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews, 2, 661-670.
33. Liberman A. M., Cooper F. S., Shankweiler, D. P., & Studdert-Kennedy, M (1967). Perception of the speech code. Psychological Review, 74, 431-461.
34. Ploog, D. (2002). Is the neural basis of vocalisation different in non-human primates and Homo sapiens? In T. J. Crow (Ed.) The speciation of modern Homo Sapiens (pp. 121-135). Oxford, UK: Oxford University Press.
35. Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. Science, 297, 846-848.
36. Semaw, S. P., Renne, P., Harris, J. W. K., Feibel, C. S., Bernor, R. L., Fessweha, N., et al. (1997). 2.5-million-year-old stone tools from Gona, Ethiopia. Nature, 385, 333-336.
37. Wood, B. (2002). Hominin revelations from Chad. Nature, 418, 134-135.
38. Ambrose, S. H. (2001). Paleolithic technology and human evolution. Science, 291, 1748-1752.
39. Walter, R. C., Buffler, R. T., Bruggemann, J. H., Guillaume, M. M. M., Berhe, S. M., Negassi, B., et al. (2001). Early human occupation of the Red Sea coast of Eritrea during the last interglacial. Nature, 405, 65-69.
40. Wood, B., & Collard, M. (1999). The human genus. Science, 284, 65-71.
41. Janis, C. (1993). Victors by default: The mammalian succession. In S. J. Gould (Ed.), The book of life (pp. 169-217). New York: W. W. Norton.
42. Suc, J.-P., Bertini, A., Leroy, S. A. G., & Suballyova, D. (1997). Towards the lowering of the Pliocene/Pleistocene boundary to the Gauss-Matuyama reversal. Quaternary International, 40, 37-42.
43. Foley, R. (1987). Another unique species: patterns in human evolutionary ecology. Harlow: Longman Scientific and Technical.
44. Tooby, J., & DeVore, I. (1987). The reconstruction of hominin evolution through strategic modeling. In W. G. Kinzey (Ed.), The evolution of human behavior: Primate models (pp. 183-237). Albany, NY: SUNY Press.
45. Pinker, S. (2003). Language as an adaptation to the cognitive niche. In M. H. Christiansen, & S. Kirby (Eds.), Language evolution (pp. 16-37). Oxford: Oxford University Press.
46. Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory, pp. 381-403. New York: Academic Press.
47. Tulving, E. (1983). Elements of episodic memory. London: Oxford University Press.
48. Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. Neurobiology of Learning and Memory, 82, 171-177.

49. Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. Genetic, Social, and General Psychology Monographs, 123, 133-167.
50. Suddendorf, T., & Corballis, M. C. (in press). The evolution of foresight: What is mental time travel and is it unique to humans? Behavioural & Brain Sciences.
51. Loftus, E. F. & Loftus, G. R. 1980. On the permanence of stored information in the human brain. American Psychologist, 35, 409-420.
52. Loftus, E. & Ketcham, K. (1994). The myth of repressed memory. New York: St. Martin's Press.
53. Roediger, H. L. & McDermott, K. B. (1995). Creating false memories - remembering words not presented in lists. Journal of Experimental Psychology: Learning Memory and Cognition, 21, 803-814.
54. Cheney, D. L., & Seyfarth, R. M. (1990). How monkeys see the world. Chicago: University of Chicago Press.
55. Goodall, J. (1986). The chimpanzees of Gombe: Patterns of behavior. Cambridge, MA: Harvard University Press.
56. Deacon, T. (1997). The symbolic species: The coevolution of language and the brain. New York: Norton.
57. Anderson, J. M. (1971). The grammar of case: Towards a localist theory. Cambridge: Cambridge University Press.
58. Cook, W. A. (1989). Case grammar theory. Washington, DC: Georgetown University Press.
59. Mandler, J. (1996). Preverbal representation and language. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), Language and space (pp. 365-384). Cambridge, MA: MIT Press.
60. Bloom, P., Peterson, M. A., Nadel, L., & Garrett, M. F. (Eds.). (1996). Language and space. Cambridge, MA: MIT Press.
61. O'Keefe, J., & Nadel, L. (1978). The hippocampus as a cognitive map. Oxford: Clarendon Press.
62. Wheeler, M. A., Stuss, D. T. & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. Psychological Bulletin, 121(3): 331-354.
63. Frisk, V. & Milner, B. (1990). The role of the left hippocampal region in the acquisition and retention of story content. Neuropsychologia, 28, 349-359.
64. Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N. F., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: evidence for a distributed model of language breakdown in aphasic patients and neurologically intact individuals. Psychological Review, 108, 759-788.
65. Karnath, H.-O., Ferber, S., & Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. Nature, 411, 950-953.
66. Rogers, L. J. & Andrew, R. J. (2002). Comparative vertebrate lateralization. Cambridge: Cambridge University Press, 2002.
67. Bramble, D. M., & Lieberman, D. E. (2004). Endurance running and the evolution of Homo. Science, 432, 345-352.
68. Swisher, C. C., III, Curtis, G. H., Jacob, A. C., Getty, A. G., Suprojo, A., & Widiasmoro. (1994). Age of the earliest known hominins in Java, Indonesia. Science, 263, 1118-1121.
69. Burling, R. (2005). The talking ape. New York: Oxford University Press.
70. Browman, C.P., & Goldstein, L.F. (1995). Dynamics and articulatory phonology. In T. van Gelder & R. F. Port (Eds.), Mind as motion (pp. 175-193). Cambridge, MA: MIT Press.
71. Studdert-Kennedy, M. (1998). The particulate origins of language generativity: From syllable to gesture. In J. R Hurford, M. Studdert-Kennedy, & C. Knight (Eds), Approaches to the evolution of language (pp. 169-176). Cambridge, UK: Cambridge University Press.
72. Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. Journal of Language and Social Psychology, 19, 163-194.
73. Goldin-Meadow S., & McNeill D. (1999). The role of gesture and mimetic representation in making language the province of speech. In M. C. Corballis & S. E. G. Lea (Eds.) The descent of mind (pp. 155-172). Oxford, UK: Oxford University Press.
74. McNeill, D. (1985). 'So you think gestures are nonverbal?' Psychological Review, 92, 350-371.
75. Gentilucci, M., Benuzzi, F., Gangitano, M. & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. Journal of Neurophysiology, 86, 1685-1699.
76. Gentilucci, M., Santunione, P., Roy, A. C. & Stefanini (2004). Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. European Journal of Neuroscience, 19, 190-202.
77. Gentilucci, M. (2003). Grasp observation influences speech production. European Journal of Neuroscience, 17, 179-184.
78. Gentilucci, M. & Corballis, M. C. 2006. From manual gesture to speech: A gradual transition. Neuroscience & Biobehavioral Reviews (in press).
79. McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746-748.
80. Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: The neural substrates of visible speech. Journal of Cognitive Neuroscience, 15, 57-70.
81. Burling, R. (1999). Motivation, conventionalization, and arbitrariness in the origin of language. In B.J. King (Ed.), The origins of language: What nonhuman primates can tell us (pp. 307-350). Santa Fe, NM: School of American Research Press.
82. Petrides, M., Cadoret, G., & Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature, 435, 1325-1328.
83. Kay, R. F., Cartmill, M., & Barlow, M. (1998). The hypoglossal canal and the origin of human vocal behavior. Proceedings of the National Academy of Sciences (USA), 95, 5417-5419.
84. DeGusta, D., Gilbert, W. H., & Turner, S. P. (1999). Hypoglossal canal size and hominin speech. Proceedings of the National Academy of Sciences, 96, 1800-1804.
85. MacLarnon, A. & Hewitt, G. (1999). The evolution of human speech: The role of enhanced breathing control. American Journal of Physical Anthropology, 109, 341-363.
86. MacLarnon, A., & Hewitt, G. (2004). Increased breathing control: Another factor in the evolution of human language. Evolutionary Anthropology, 13, 181-197.
87. Lieberman, D. E. (1998). Sphenoid shortening and the evolution of modern cranial shape. Nature, 393, 158-162.
88. Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the new-born, adult human, Neanderthal man, and the chimpanzee. American Anthropologist, 74, 287-307.
89. Gibson, K. R., & Jessee, S. (1999). Language evolution and expansions of multiple neurological processing areas. In B. J. King (Ed.), The origins of language: What nonhuman primates can tell us. Santa Fe, NM: School of American Research Press.
90. Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. Yearbook of Physical Anthropology, 45, 36-62.
91. Vleck, E. (1970). Etude comparative onto-phylogénétique de l'enfant du Pech-de-L'Azé par rapport à d'autres enfants néanderthaliens. In D. Feremback (Ed.), L'enfant Pech-de-L'Azé (pp. 149-186). Paris: Masson.

92. Lieberman, D. E., McBratney, B. M., & Krovitz, G. (2002). The evolution and development of cranial form in Homo sapiens. Proceedings of the National Academy of Sciences, 99, 1134-1139.
93. Vargha-Khadem, F., Watkins, K.E., Alcock, K.J., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. Proceedings of the National Academy of Sciences, USA, 92, 930-933.
94. Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A.P., Pembrey, M.E., (1998). Localisation of a gene implicated in a severe speech and language disorder. Nature Genetics, 18, 168-170.
95. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco A. P. (2001). A novel forkhead-domain gene is mutated in a severe speech and language disorder. Nature, 413, 519-523.
96. Gopnik, M. (1990). Feature-blind grammar and dysphasia. Nature, 344, 715.
97. Watkins, K.E., Dronkers, N.F., & Vargha-Khadem F. (2002). Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. Brain, 125, 452-464.
98. Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. Nature Neuroscience, 6, 1230-1237.
99. Watkins, K.E., Vargha-Khadem, F., Ashburner, J., Passingham, R.E., Connelly, A., Friston, K.J., et al. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. Brain, 125, 465-478.
100. Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. Nature, 418, 869-871.
101. Marcus, G. F., Fisher, S. E. (2003). FOXP2 in focus: what can genes tell us about speech and language? Trends in Cognitive Science, 7, 257-262.
102. Shu, W.G., Yang, H.H., Zhang, L.L., Lu, M.M., & Morrisey, E.E. (2001). Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. Journal of Biological Chemistry, 276, 27488-27497.
103. Knight, A., Underhill, P.A., Mortensen, H.M., Zhivotovsky, L.A., Lin, A.A., Henn, B.M., et al. (2000). African Y chromosome and mtDNA divergence provides insight into the history of click languages. Current Biology, 13, 464-473.
104. Mellars, P. (2006). Going east: New genetic and archaeological perspectives on the modern human colonization of Eurasia. Science, 313, 796-800.
105. Oppenheimer, S. (2003). Out of Eden: The peopling of the world. London: Constable.
106. Olivieri, A., Achilli, A., Pala, M., Battaglia, V., Fornarino, S., Al-Zaherym N., et al. (2006). The mtDNA legacy of the Levantine Early Upper Palaeolithic in Africa. Science, 314, 1767-1770.
107. Russell, B.A., Cerny, F.J., & Stathopoulos, E.T. (1998). Effects of varied vocal intensity on ventilation and energy expenditure in women and men. Journal of Speech, Language, and Hearing Research, 41, 239-248.
108. Konner, M. (1982). The tangled wing: biological constraints on the human spirit. New York: Harper.
109. Boutla, M., Supalla, T., Newport, E. L., & Bavelier, D. (2004). Short-term memory span: Insights from sign language. Nature Neuroscience, 7, 997-1002.
110. Mellars, P. A. & Stringer, C. B. (Eds.) (1989). The human revolution: Behavioural and biological perspectives on the origins of modern humans. Edinburgh: Edinburgh University Press.
111. McBrearty, S. & Brooks, A. S. (2000). The revolution that wasn't: A new interpretation of the origin of modern human behavior. Journal of Human Evolution, 39, 453-563.
112. Corballis, M. C. (2004). The origins of modernity: Was autonomous speech the critical factor? Psychological Review, 111, 543-522.
113. Pepperberg, I. M. (2002). In search of King Solomon's ring: Cognitive and communicative studies of Gray parrots (Psittacus erithacus). Brain, Behavior, and Evolution, 59, 54-67.
114. Kaminsky, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping." Science, 304, 1682-1683.
115. Christiansen, M. H., & Kirby, S. (2003). Language evolution: The hardest problem in science? In M. H. Christiansen, & S. Kirby (Eds.), Language evolution (pp. 1-15). Oxford: Oxford University Press.