

# 1 Origins of Language

---

ANDREW CARSTAIRS-MCCARTHY

## 1 Introduction

Among the inhabitants of some African forests about eight million years ago were ape-like creatures including the common ancestors of chimpanzees and humans. Visualizing what these creatures were probably like is easy enough; one conjures up an image of something resembling a modern gorilla, living substantially in trees and walking on all four limbs when on the ground, and with a vocal communication system limited to perhaps twenty or thirty calls, like a chimpanzee's. But what about our ancestors' appearance and behavior two million years ago? By that stage they were a separate species from the ancestors of chimpanzees, but were not yet *homo sapiens*. How did these creatures live, and in particular what sort of language did they have? Visualizing these more recent creatures is harder. One feels that they must have been more like us, and in particular that their vocal communication system must have been more sophisticated than that of their ancestors six million years earlier. But how much more sophisticated? Which characteristics of modern human language did this communication system now possess, and which did it still lack?

There is something eerie and yet fascinating about these intermediate ancestors. This fascination underlies innumerable science fiction stories as well as the perennial interest in rumors that such creatures may still exist, in some remote Himalayan valley perhaps. To many nonlinguists, therefore, it seems self-evident that research on the linguistic abilities of such intermediate ancestors (that is, research on the origins and evolution of human language) should be a high priority in linguistics. Yet it is not. As a research topic, language evolution is only now beginning to regain respectability, after more than a century of neglect. In the remainder of this section I will say something about the reasons for this neglect before turning in sections 2–5 to the evidence recently brought to bear by anthropologists, geneticists, primatologists and neurobiologists, many

of whom have for decades been more adventurous than linguists in this area. Then in section 6, I will discuss the kinds of contribution which some linguists also are now beginning to offer.

Many religions provide an account of the origin of language. According to the Judeo-Christian tradition, God gave to Adam in the Garden of Eden dominion over all the animals, and Adam's first exercise of this dominion consisted in naming them. The fact that there are now many languages rather than just one is explained in the story of the Tower of Babel: linguistic diversity is a punishment for human arrogance. So long as that sort of account was generally accepted, the origin of language was not a puzzle. But when secular explanations for natural phenomena began to be sought to supplement or replace religious ones, it was inevitable that a secular explanation was sought for the origin of language too.

The fact that the origin of language must predate recorded history did not inhibit eighteenth-century thinkers such as Rousseau, Condillac, and Herder, who were confident that simply by applying one's mind to the situation in which languageless humans would find themselves one could arrive at worthwhile conclusions about how language must have arisen. Unfortunately there was no consensus among these conclusions, and in the nineteenth century they came to seem increasingly feeble and speculative by contrast with the far-reaching yet convincing results attainable in historical and comparative linguistics (see chapter 5). At its foundation in 1866, therefore, the Linguistic Society of Paris chose to emphasize its seriousness as a scholarly body by including in its statutes a ban on the presentation of any papers concerning the origin of language. Most linguists still support this ban, in the sense that they believe that any inquiry into the origin of language must inevitably be so speculative as to be worthless.

Since the 1960s, the theory of grammar has come to be dominated by the ideas of Noam Chomsky, for whom the central question of linguistics is the nature of the innate biological endowment which enables humans to acquire a language so rapidly and efficiently in the first year of life (see chapter 19). From this viewpoint, it seems natural to regard the origin of language as a matter of evolutionary biology: how did this innate linguistic endowment evolve in humans, and what are its counterparts (if any) in other primates? But Chomsky has explicitly discouraged interest in language evolution, and has even suggested that language is so different from most other animal characteristics that it may be a product of physical or chemical processes rather than biological ones (1988: 167, 1991: 50). The paradoxical result is that, while Chomskyan linguists endeavor to explain characteristics of individual languages by reference to an innate linguistic endowment (or Universal Grammar), they are generally reluctant to pursue their inquiry one stage further, to the issue of how and why this innate endowment has acquired the particular characteristics that it has. To be sure, there are exceptions (e.g. Newmeyer 1991, Pinker and Bloom 1990, Pinker 1994). Nevertheless, Chomsky's influence means that linguists' reluctance to tackle this area is eroding only slowly.

In view of what has been said, it is not surprising that there is a shortage of introductory surveys of this topic from a linguistic point of view; but Aitchison (1996) can be recommended, as well as part II of W. Foley (1997). Hurford et al. (1998) is an up-to-date collection of contributions from a variety of disciplines.

## 2 Evidence from Anthropology and Archeology

Anthropology is concerned not only with human culture but also with humans as organisms in a biological sense, including their evolutionary development. (On human evolution in general, see e.g. R. Foley (1995) and Mithen (1996).) Language is both a cultural phenomenon and also the most salient distinguishing characteristic of modern *homo sapiens* as a species. The question of how and why humans acquired language therefore interests both cultural and biological anthropologists. So what light can anthropology shed on these questions?

The earliest direct evidence of language in the form of writing is no more than about 5,000 years old (see chapter 3). It is therefore much too recent to shed any light on the origin of spoken language, and we must resort to indirect evidence. Unfortunately the available evidence is doubly indirect. The vocal apparatus (tongue, lips, and larynx) of early humans would tell us much if we could examine it directly; but, being soft tissue, it does not survive, and for information about it we have to rely on what we can glean from bones, particularly skulls. Alongside such evidence we have tools and other artefacts, as well as traces of human habitation such as discarded animal bones; but, again, what is available to us is skewed by the fact that stone survives better than bone and much better than materials such as wood or hide. In view of this, the only relatively firm dates which anthropology can provide are two terminuses, one after which we can be sure that language in its fully modern form did exist and one before which we can be sure that it did not. For the long period in between, the anthropological evidence is tantalizing but frustratingly equivocal; there are no uncontroversial counterparts in the fossil record for specific stages in linguistic evolution.

We can be reasonably confident that modern-style spoken language evolved only once. This is not logically necessary. It is conceivable that something with the communicative and cognitive functions of language, and using speech as its medium, could have evolved independently more than once, just as the eye has evolved independently more than once in the animal kingdom. However, if that had happened we would expect to find evidence of it today, just as the eyes of octopuses, mammals, and insects reveal by their structure that they have no common ancestor. Yet no such evidence exists. For all their diversity, all existing languages display certain fundamental common properties of grammar, meaning, and sound, which is why Chomsky feels justified in claiming

that, to a visitor from another planet, it might seem that there really is only one human language. Moreover, a child who is removed from her parents' speech community at a young age can acquire natively any language whatever, irrespective of what her parents speak; no child is born with a biological bias in favor of one language or type of language. This means that language of a fully modern kind must have evolved before any contemporary human group became geographically separated from the rest of the human race (separated, that is, until the invention of modern means of transport). The first such clearcut separation seems to have occurred with the earliest settlement of Australia by *homo sapiens*. Archeological evidence suggests that that event took place at least 40,000 years and perhaps as long as 60,000 or more years ago. We can therefore take this as a firm *terminus ante quem* for the evolution of a form of language which is fully modern in a biological sense.

As for a *terminus post quem*, it is clear that spoken language with more or less modern articulatory and acoustic characteristics presupposes something like a modern vocal tract. But how are we to interpret "more or less" and "something like"? One thing is clear: the acoustic properties of many human speech sounds, particularly vowels, depend on the characteristically human L-shaped vocal tract, with an oral cavity at right angles to the pharynx (see chapter 7) and with the larynx relatively low in the neck. This shape is characteristically human because in nearly all other mammals, and even in human babies during the first few months of life, the larynx is high enough for the epiglottis to engage with the soft palate so as to form a self-contained airway from the nose to the lungs, smoothly curved rather than L-shaped, and quite separate from the tube which leads from the mouth to the stomach. Having these two distinct tubes enables nearly all other mammals, as well as newborn human babies, to breathe while swallowing. The adult human characteristic of a pharynx through which both air and food must pass, on the other hand, is a vital contributor to the acoustic characteristics structure of speech sounds. So when did this L-shaped vocal tract develop?

Lieberman (1984, see Lieberman and Crelin 1971) has claimed that even in Neanderthals, who did not become extinct until about 35,000 years ago, the larynx was positioned so high in the neck as to prevent the production of the full modern range of vowel sounds. He suggests that this linguistic disadvantage may have been a factor in the Neanderthals' demise. But his argument rests on an interpretation of fossil cranial anatomy which has generally been rejected by anthropologists (Trinkaus and Shipman 1993, Aiello and Dean 1990). An alternative view is that the L-shaped vocal tract is a byproduct of bipedalism, which favored a reorientation of the head in relation to the spine and hence a shortening of the base of the skull, so that the larynx had to be squeezed downward into the neck (DuBrul 1958, Aiello 1996b). The question then arises: when did our ancestors become bipedal? The general consensus among anthropologists is: very early. Evidence includes fossil footprints at Laetoli in Tanzania, about 3.5 million years ago, and the skeleton of *australopithecus afarensis* nicknamed "Lucy," dating from over 3 million years ago. So, if bipedalism was the

main factor contributing to the lowering of the larynx, the L-shaped vocal tract must have emerged relatively early too.

This conflicts with an opinion widespread among language origin researchers, namely that the lowering of the larynx (with its concomitant increased risk of choking) was a consequence of the evolution of more sophisticated language, not a precursor of it (a “preadaptation” for it, in Darwinian terminology). But this predominant opinion, it may be argued, is to some extent a hangover from the “brain-first” view of human evolution in general – the view that the superior intelligence of humans evolved in advance of substantial anatomical changes, broadly speaking. This view was popular when Piltown Man, with its human-like skull and ape-like jaw, was still thought to be genuine, but is now generally rejected in the face of evidence for the small size of australopithecine and early human skulls.

Mention of skulls raises the possibility of drawing conclusions about language from hominid brains. (I use the term *hominid* to mean “(belonging to) a creature of the genus *australopithecus* or the genus *homo*.”) Brain size tells us nothing specific (though we will revert to it in section 6). But what of brain structure? If it could be shown that an area of the modern human brain uniquely associated with language was present in the brains of hominids at a particular date, it would seem reasonable to conclude that those hominids possessed language. But this line of reasoning encounters three problems. First, since brains themselves do not fossilize, determining their structure depends on the interpretation of ridges and grooves on the inside of skulls, or rather of their counterparts on “endocasts” made from skulls. The region generally regarded as most closely associated with grammar and with speech articulation in modern humans is Broca’s area; but identifying an area corresponding to Broca’s area in hominid fossils has turned out to be highly controversial (Falk 1992). Second, no area of the human brain, even Broca’s area, seems to be associated with language and nothing else. Third, Broca’s area seems to have little or nothing to do with vocalization in monkeys, so even if it can be established that a counterpart of Broca’s area exists in a certain hominid, its function in that hominid may not be linguistic. The implications of “brain-language coevolution,” as Deacon (1997) calls it, are still frustratingly indeterminate.

Some scholars have connected language with the evolution of “handedness,” which is much more strongly developed in humans than in other animals (Bradshaw and Rogers 1992, Corballis 1991). In most people the right hand is the dominant hand, controlled from the left side of the brain where the language areas are usually located. It is tempting to see this shared location as more than mere coincidence. If so, linguistic conclusions might perhaps be drawn from ingenious tests that have been carried out on fossil stone tools, to determine whether the people who made them were or were not predominantly right-handed. On the other hand, the correlation between language and handedness is far from exact: left-handedness neither entails nor is entailed by right-brain dominance for language. Also, even if evidence of a strong preponderance of right-handers in some group of hominids is taken as firm evidence

of linguistic capacity, it furnishes no details about the nature of that linguistic capacity.

Let us turn from biology to culture. Common sense would suggest that a relatively sudden jump in the sophistication of human linguistic behavior, if it occurred, should leave immediate traces in the archeological record in the shape of a sudden jump in the sophistication of preserved artefacts (tools, ornaments, and artwork). So does any such jump in sophistication occur, and when? There is indeed a big increase in the variety and quality of tools found in Europe and Africa around 40,000 years ago, followed by the famous cave paintings of Lascaux and elsewhere from about 30,000 years ago. But this is inconveniently late as a date for the emergence of fully modern language, in that it is contemporary with or even more recent than the latest plausible date for the settlement of Australia. That has not discouraged some scholars from using this kind of evidence to argue that language evolved “late”; but on examination it generally turns out that what these scholars mean by “language” is not what linguists mean by it, but rather the self-conscious use of symbols (Noble and Davidson 1996). Moreover, there is scattered but intriguing evidence of “cultural” behavior thousands of years earlier, such as burial pits, incised bones and the use of red ochre pigment for body decoration. The implications of this for language are unclear, but it may be significant that some of the dates involved are not far removed from a milestone indicated by genetic evidence, to which we now turn.

### 3 Genetic Evidence

Since the late 1970s, molecular genetics has opened up entirely new techniques for assessing the relationship of humans to each other and to other primates. (It is genetic evidence which tells us that we are separated by only about five million years from the ancestor which we share with the chimpanzees.) Since the 1950s it has been known that the information which differentiates an individual genetically from all other individuals (except a possible identical twin) is carried by DNA (deoxyribonucleic acid) in chromosomes located in every cell in the body. Geneticists can now compare individuals and groups in terms of how much of their DNA is shared. Moreover, they can do this not only in respect of the DNA in the cell’s nucleus, which is inherited from both parents, but also in respect of the DNA in the cell’s mitochondria – some of the so-called “organelles” which the cell contains in addition to its nucleus. What is important about mitochondrial DNA is that it is inherited from the mother alone. It follows that the only reason that there can be for any difference between two people’s mitochondrial DNA is inaccurate inheritance due to mutation; for, without this inaccuracy, both of them would have exactly the same mitochondrial DNA as their most recent shared ancestor in the female line. So, assuming that mutation in DNA occurs at a constant rate, the

extent of difference between two people's DNA is an indication of the number of generations which separate them from the most recent woman from whom both are descended through her daughters, her daughters' daughters, and so on.

Cann et al. (1987) used this technique to try to locate in time and space the most recent woman from whom all living humans are descended in the female line. With the help of elaborate statistical techniques, they argued that this woman lived roughly 200,000 years ago in Africa, hence the nickname "African Eve." Both the African location and the date corresponded quite closely to the "out-of-Africa" scenario for early *homo sapiens* proposed on independent grounds by some archeologists, so the two theories provided mutual support. The nickname "Eve" is convenient but unfortunate, because it suggests that, apart from Eve's male partner or partners, none of her contemporaries has any descendants alive today. That is a fallacy; all one can say is that anyone alive today who is descended from a female contemporary of Eve must be linked to that woman through at least one male ancestor. However, the argument of Cann and her colleagues does suggest that there was a population bottleneck relatively recently in human prehistory, such that most of the humans alive around 200,000 years ago, scattered over large areas of Africa, Europe, and Asia, have indeed left no surviving descendants. Why should this be?

Many scholars have been tempted to suggest that what was special about Eve's community – the characteristic which enabled their descendants to outperform other humans and which discouraged interbreeding with them – must have been superior linguistic abilities, presumably newly acquired. This is only a guess, however. Cann herself has more recently mentioned one of many alternative possibilities: infectious disease (Cann et al. 1994). But the possible link with language evolution has been popularized by Cavalli-Sforza and Cavalli-Sforza (1995) and by Ruhlen (1994), whose supposed reconstructions of Proto-World vocabulary might, if genuine, be roughly contemporary with Eve. An equivocation on "mother tongue" underlies this view, however. Even supposing it were possible to reconstruct the most recent language from which all contemporary languages are descended, it would be a remarkable coincidence if that ancestral language (the "mother tongue" in a historical linguistic sense) were also the first linguistic variety with fully modern characteristics (the "mother tongue" in a biological sense). So, once again, we are faced with evidence which, though tantalizing, does not point to any firm conclusion.

## 4 Primatological Evidence

No living primate apart from man is equipped to speak. However, three areas of current research on primates may shed light on language evolution.

These involve primate vocal call systems, primate cognitive abilities (particularly their knowledge of social relationships), and the results of experiments involving teaching sign language and artificial signaling systems to apes.

### **4.1 *Vocal call systems***

Until a few decades ago, it was generally thought that the calls uttered by all animals, including monkeys and apes, were exclusively reflections of physical or emotional states such as pain, fear, hunger, or lust. In this respect, the portion of the human vocal repertoire which primate call systems seemed to resemble most closely was the portion consisting of involuntary sounds such as cries of pain, laughter, or sobbing. No linguists have been reluctant to contemplate an evolutionary link between these cries and primate vocalizations. But primate “vocabularies” were thought to lack a central element of human vocabularies: referential calls identifiable with specific objects or classes of objects in the external world. Given that assumption, it was easy to dismiss animal calls systems as irrelevant to human language. However, students of animal behavior were becoming increasingly uncomfortable with this assumption, and Cheney and Seyfarth (1990) developed a particularly elegant and convincing way of testing it systematically. (On animal communication generally, see Hauser 1996; on the calls of chimpanzees in the wild, see Goodall 1986.)

In the 1970s and 1980s, Cheney and Seyfarth spent years investigating the behavior of vervet monkeys in their native habitat, the Amboseli National Park of Kenya. These small monkeys utter distinct warning calls for different types of predator, notably leopards, snakes, and eagles, for which different types of evasive action are appropriate: they run up trees to escape leopards, peer at the ground around them to avoid snakes, and hide in bushes to evade eagles. This kind of apparent referentiality had been noticed before, not just among vervets; but such awareness had not shaken the general conviction among both zoologists and linguists that animal cries were basically emotional or affective in content rather than referential. In crude terms, a vervet’s eagle call would be interpreted as linked not to something in the outside world (“There’s an eagle!”) but rather to its internal state (“I am experiencing eagle-fear!” or “I feel an urge to hide in bushes!”). To be sure, if one vervet uttered the eagle call, others might take evasive action too; but this could only be because these others saw the eagle for themselves and hence experienced the same emotion (it was thought).

Cheney and Seyfarth showed this interpretation to be incorrect by a crucial experiment. They made recordings of predator warning calls and played them back from hidden loudspeakers in the absence of the relevant predators. If the traditional interpretation of the warning calls was correct, the vervets would be predicted to take no evasive action in response to these bogus calls. They might look around for the relevant predator but, failing to see one, they would



not experience the relevant fear reaction and so would do nothing. However, what Cheney and Seyfarth found was that the vervets reacted to the bogus calls just as if they were genuine, by taking the appropriate evasive action. The call itself was the trigger to act, not the emotion or physical state engendered by the sight of a predator. Warning calls therefore really do contain referential information about the environment, on which vervets can act appropriately. To this admittedly limited extent, therefore, they resemble words of a human language.

A second respect in which human language differs from animal cries, it used to be thought, is that only human language can be unreliable. If an animal cry is an automatic response to an emotional or physical stimulus, its reliability is in some sense guaranteed. Humans, on the other hand, can tell lies or make mistakes. But Cheney and Seyfarth showed that in this respect too the gap between vervet monkeys' calls and human language is less than was once thought. Vervets' use of their warning calls is not entirely innately determined; for example, young vervets will often utter the eagle call even when they have seen something in the sky which is not an eagle or not even a bird at all, such as a falling leaf. And adult vervets react differently to young vervets' calls too. Instead of taking immediate evasive action, as they would if they had heard an adult call, they first check for themselves whether the relevant predator is present and, if not, ignore the call. It seems to be through observing when its calls are acted upon and when they are ignored that a young vervet refines its innate repertoire of vocal reactions into accurate warnings deployed according to the conventions of the adult community.

These observations show that, for vervets, calls have a content which is independent of their own physical or emotional state. Cheney and Seyfarth were also able to show that, in judging the reliability of a call that it hears, a vervet goes beyond merely identifying the caller. It is clear that vervets can distinguish individual "voices," because when a young vervet utters a cry of distress the adults in earshot will look towards that individual's mother, as if expecting her to respond. Cheney and Seyfarth compared reactions to recordings of different voices uttering a variety of calls. In the absence of a genuine eagle danger, hearers will become habituated to and hence ignore recorded eagle alarms in the voice of vervet A, but will still react to alarms in the voice of vervet B. But, even when so habituated to vervet A, they will not ignore a recording of vervet A uttering a call of a different kind (say one of the repertoire of calls relating to individual or group interactions). Vervets can evidently distinguish, in respect of another vervet, those topics on which it is a reliable witness from those on which it is unreliable.

To be sure, the vervet call system has no grammatical organization remotely resembling that of human language, and the same is true of all other primates' call systems. Nevertheless, the observations of Cheney, Seyfarth, and others tend to show that the differences between primate call systems and human language are not so great as was once thought, and hence weaken the case for denying any evolutionary connection between them.

## 4.2 *Cognitive abilities*

Long-term observations of primate groups in the wild, such as those of Goodall and Cheney and Seyfarth mentioned in section 4.1, show that primates know many more details about themselves, their conspecifics and their environment than was previously suspected. In particular, they can distinguish kin from nonkin, and by remembering who has done what to whom they can distinguish allies from enemies. This is relevant to language inasmuch as a fundamental characteristic of language is the ability to represent grammatically the roles of participants in a situation (Bickerton 1990, 1995). For example, the sentence *John gave Mary a banana* represents a situation in which John is the agent, Mary is the goal and the banana is the patient or “theme” in relation to an act of giving. In the terminology of semantics, such a set of relationships between participants in a situation is called a “thematic structure” or “argument structure” (see chapter 12). Higher primates do not produce sentences, but they certainly have mental representations of thematic structures of the kind which underlie sentences. To that extent they have evolved to a stage of cognitive readiness for language.

One of the Rubicons which have been claimed to separate humans from other animals is that, whereas other animals may possess “procedural” knowledge (“knowledge-how”), only humans have access to “propositional” knowledge (“knowledge-that”). (In similar vein, Donald (1991) distinguishes between “episodic,” “mimetic,” and “mythic” culture, among which only “episodic” culture is available to nonhumans.) If this is correct, it is tempting to see propositional knowledge as a prerequisite for language. In assessing whether it is correct, however, one immediately encounters a risk of circularity. If “propositional knowledge” means simply “knowledge of a kind which can only be represented in sentence form,” then it is not surprising that propositional knowledge should be restricted to sentence-users, that is, to humans; but then to say that animals lack it is to say no more than that animals lack language. On the other hand, if “propositional knowledge” is defined so as to make it logically independent of language, such as in terms of thematic structure, it is by no means so clear that this Rubicon exists.

At least two further considerations support the idea that primates have access to “knowledge-that.” One is the extent to which, in the admittedly artificial conditions of the laboratory, chimpanzees can acquire and display awareness of abstract concepts such as “same” and “different” and apply them by reference to a range of criteria such as color and size (Premack 1976). The other is the fact that primates can apparently indulge in deception, or display what has been called “Machiavellian intelligence” (Byrne and Whiten 1988, Sommer 1992). In interpreting “Machiavellian” behavior it is of course necessary to guard against overenthusiastic ascription of human personality traits to animals. Nevertheless, this behavior suggests that primates are capable of conceiving of situations which do not exist, that is to think in an abstract “propositional”

fashion, and hence reinforces the worthwhileness of looking for precursors of language in other species.

Social relationships among primates are both more complex and less stereotyped than among other mammals, and it has been suggested that social factors may outweigh communicative ones in fostering language evolution. Dunbar (1996) and others have drawn attention to the relationship between group size, brain size, and social grooming in various primate species. Grooming is important in fostering group cohesion; on the other hand, time devoted to grooming increases exponentially as group sizes increase, thereby reducing the time available for other essential tasks such as food gathering. Dunbar suggests that language provided a way out of this dilemma: it is a form of vocal grooming, with the advantage that by means of language one can groom many other individuals at once. Traces of this original function can be observed in the extent to which, even today, language is used for gossip and for cementing social relationships rather than for the more abstract representational and information-conveying purposes which tend to interest grammatical theorists and philosophers.

### ***4.3 Sign language experiments***

Apes do not have vocal tracts fitted for speech, but their arms and hands are physically quite capable of forming the signs of deaf languages such as American Sign Language (ASL). In the 1970s great excitement was generated by experiments which purported to show that chimpanzees could learn ASL, so that language could no longer be regarded as a uniquely human attribute (Terrace 1979, Gardner et al. 1989). Linguists in general hotly denied that the sign sequences produced by chimpanzees such as Washoe and Nim could be regarded as genuine syntactic combinations or complex words, pointing to the fact that the chimpanzees' sign sequences never reached the variety and complexity of those of fluent human ASL signers. The chimpanzees' supporters, on the other hand, argued that the kinds of sign combination which chimpanzees produced were quite similar to the word combinations which human babies produce at the "two-word" or "telegraphic" stage of language acquisition, so that, if what the chimpanzees did was not a manifestation of language, one could not call babies' "telegraphic" speech a manifestation of language either. (We will return to this implication in section 6.) In the present context the issue is not whether the chimpanzees' and other apes' signing behavior can properly be called linguistic (which is in any case largely a sterile point of terminology), but whether this behavior sheds any light on language evolution.

One effect of the ape language experiments was to give new life to the old idea that language in humans may have originated in gesture, and only later been transferred to the vocal channel (Armstrong et al. 1995). Just as apes can sign without a human vocal tract, so could our australopithecine ancestors

have communicated by sign before their vocal tracts had become capable of modern-style speech, perhaps. One of the attractions of this proposal has always been that it seems to provide a solution to the problem of how humans originally learned to handle the arbitrary relationship between words and meanings. The apparent solution lies in the fact that many signs in ASL and other sign languages are motivated (“iconic”) rather than arbitrary (“symbolic”); that is, they resemble or recall in some way their referents in the outside world, while many other signs were once more clearly motivated than they are now. The proportion of sign language vocabularies which is iconic is far greater than the proportion of iconic (onomatopoeic) words in spoken language vocabularies. These motivated manual signs could have constituted a scaffolding, so to speak, to assist the more difficult task of mastering arbitrary signs, whether manual or vocal. But the attraction of this reasoning disappears as soon as one recalls that vervet monkeys’ call vocabulary is just as symbolic as most words of human language. Vervets’ eagle, leopard, and snake calls do not in any way resemble or sound like eagles, leopards, or snakes. So, even if one regards the use of symbolic signs as a communicative Rubicon, it is a Rubicon which has been crossed by any nonhuman species with a clearly referential call vocabulary, and was almost certainly crossed by our primate ancestors long before the appearance of hominids.

More relevant to language evolution, perhaps, is what can be gleaned from observation of the bonobo (or pygmy chimpanzee) Kanzi (Savage-Rumbaugh et al. 1993, Savage-Rumbaugh and Lewin 1994). Savage-Rumbaugh set out to train Kanzi’s mother in both sign language and the use of a keyboard of arbitrary wordsigns or “lexigrams,” while the infant Kanzi was left to play and watch what was going on unmolested. The mother turned out to be an unpromising pupil; Kanzi, on the other hand, developed spontaneously a form of communication involving both manual signs and lexigrams, and also showed a surprising ability to understand spoken English – a somewhat more accurate understanding, in fact, than the two-year-old daughter of one of Savage-Rumbaugh’s colleagues, at least within a deliberately limited range of syntactic constructions.

Savage-Rumbaugh argues that Kanzi shows evidence of rule-governed use of signs and lexigrams, and one may if one wishes call this set of rules a syntax. But it seems overhasty to conclude, as Savage-Rumbaugh does, that the difference between Kanzi’s syntax and that of human languages is only in degree of complexity, not in kind. Of the two rules which Kanzi has invented rather than merely copied from human sign use, one (“lexigram precedes gesture”) clearly has no human counterpart, while the other (“action precedes action, sign order corresponding to order of performance,” as in *chase hide* or *tickle bite*) is interpretable as purely semantic or pragmatic rather than syntactic. Moreover, Kanzi’s “utterances” are nearly all too short to permit clearcut identification of human-language-like phrases or clauses. A more conservative conclusion would be that Kanzi may indeed have invented a kind of rudimentary syntax, but it cannot be straightforwardly equated with the kind of syntax

that human languages have. A task for the language evolution researcher, then, is to account for the differences between what the bonobo does and what humans do.

## 5 Neurobiological Evidence

To investigate systematically the relationship between language and the brain, one would need to carry out surgical experiments of an ethically unthinkable kind. Our knowledge has therefore to be gleaned in a relatively haphazard fashion, from the linguistic behavior of people suffering from brain damage due to accident or cerebral hemorrhage. This is less than ideal, because the extent of the damage is of course not subject to any experimental control and is determinable only indirectly, through methods such as magnetic resonance imaging (MRI), which is like an X-ray but much more detailed, and positron emission tomography (PET), which measures minute changes in bloodflow. With the patient's consent, it is also possible to test the linguistic effect of stimulating areas of brain tissue directly in the course of surgery for purposes such as the control of epilepsy (Calvin and Ojemann 1994). Not surprisingly, the literature on such research, though extensive, is somewhat confusing. However, it does suggest answers (though by no means conclusive ones) to two broad questions relevant to language evolution. The first question concerns the relative priority of the vocal and gestural channels for speech. The second concerns the extent to which syntax is an outgrowth of a general increase in the sophistication of hominids' mental representation of the world, including social relationships, and the extent to which it is an outgrowth of some more specialized development, such as better toolmaking, more accurate stone-throwing, or more fluent vocalization.

Before we consider these broad questions, it is worth emphasizing that the relationship between particular functions and particular brain locations is not clearcut and unchanging, either in the individual or in the species. Exercising one finger can increase the area of brain cortex devoted to controlling it, and in many blind people the cortex areas for finger control are larger than average. This functional plasticity is particularly evident in early infancy, so that a young child who suffers massive damage to the left brain hemisphere (where the control of language is generally located) may nevertheless acquire a considerable linguistic capacity, controlled from the right hemisphere. Indeed, without such plasticity and scope for functional overlapping it is hard to see how language could have evolved at all, because it must have involved a new role for parts of the brain which originally served other functions.

The brain region which seems most clearly implicated in regulating grammar is Broca's area, in the frontal lobe of the left hemisphere. In view of the scope for overlap in functions, it seems reasonable to predict that, if language was originally gestural, Broca's area would be relatively close to that part

of the brain which controls movement of the hands; but it is not. Control of bodily movements resides on the so-called motor strip, just in front of the central sulcus or Rolandic fissure which separates the frontal lobe from the parietal lobe. Broca's area is indeed close to the motor strip; but it is closest to that part of the strip which controls not the hands but, rather, the tongue, jaw, and lips. Moreover, a similarly located Broca's area seems to be just as relevant to the grammar of sign language, even among people deaf from birth, as it is to the grammar of spoken language (Poizner et al. 1987).

Conceivably, the region for grammatical control could have migrated, so to speak, if the predominant channel for language switched from gesture to speech. However, since the present location of Broca's area does not prevent it from playing a role in sign language, a hypothetical language area located close to the manual section of the motor strip could presumably have retained its original control over grammar even while the vocal apparatus took over from the hands. So it seems more likely that the linguistic function exercised by Broca's area has not migrated, and its present brain location reflects the fact that human language has always been predominantly vocal.

Damage to Broca's area affects grammar and also the articulation of speech much more than vocabulary. Broca's aphasics can generally produce appropriate nouns, adjectives, and verbs for what they are trying to say; it is the task of stringing them together in well-formed sentences with appropriate grammatical words (determiners, auxiliaries, and so on) which causes them trouble. A complementary kind of aphasia, involving fluent grammar but inappropriate or nonsensical vocabulary, is associated with damage elsewhere in the left hemisphere, in a region of the temporal lobe and part of the parietal lobe known as Wernicke's area. In Wernicke's aphasics the grammatical equipment to talk about the world is intact, but access to the concepts for organizing their experience of the world (insofar as one can equate concepts with items of vocabulary) is disrupted. Wernicke's aphasia is therefore problematic for any suggestion that conceptual relationships such as thematic structures (mentioned earlier) were not merely a necessary condition for the evolution of syntax but, rather, the main trigger for it. On the basis of that suggestion, one would expect lexical and grammatical disruption regularly to go hand in hand, rather than to occur independently. So, in answer to our second question, the characteristics of Wernicke's aphasia suggest that, for syntax to evolve as it has, something more specialized than just general conceptual sophistication was necessary.

Various suggestions have been made concerning this more specialized ingredient. Some scholars have appealed to the hierarchical organization of relatively complex behaviors involving tools (e.g. Greenfield 1991). Calvin (1993) has pointed out the neurobiological advances necessary for muscular control in accurate throwing, and has suggested that the relevant neural structures may have been coopted for rapid, effortless syntactic organization of words in speech. Such approaches do not, however, account for the proximity of Broca's area to that part of the motor strip which controls the mouth in particular.

One suggestion which exploits that proximity will be discussed in the next section.

## 6 Linguistic Evidence

It may seem paradoxical that the section on linguistic evidence for the origins of language has been left until last. However, as explained in section 1, linguists have been relative latecomers to this field. Their contributions can be divided into those focussing on the relationship between language and “protolanguage” and a more recent and disparate group focussing on the evolutionary rationale for particular aspects of modern grammatical organization.

### 6.1 *Protolanguage and “true” language*

Students of language contact distinguish between pidgins, which are used as second languages in situations of regular contact between people with mutually unintelligible mother tongues, and creoles, which arise when children acquire pidgins natively. The creolization process involves faster spoken delivery and the rapid appearance of new grammatical features which may be expressed unsystematically or not at all in the parent pidgin. Study of creole formation, especially among children of workers on Hawaiian sugar plantations, led Bickerton (1981) to the controversial proposal that, in environments where creoles originate, the universal human linguistic “bioprogram” reveals its characteristics most plainly, because the local speech community lacks entrenched grammatical habits which might interfere with it.

Since proposing the bioprogram hypothesis, Bickerton has turned his attention to how the bioprogram may have evolved, and to what sort of linguistic capacity may have preceded it (1990, 1995). He has suggested that what preceded it is still present and in use among humans in certain situations: in “touristese,” in the speech of people who are intoxicated or suffering from some kinds of brain damage, and especially in the “two-word” or “telegraphic” stage of infant speech already mentioned in section 4. This kind of language lacks any systematic grammar, so to understand it one must rely heavily on semantic and pragmatic cues. In particular, it lacks any systematic encoding of thematic structure of the kind which, in “true” language, allows us to distinguish reliably between agents, patients, beneficiaries, instruments, and so on (see section 4). In the English sentence *John killed a crocodile*, the identity of the agent and the patient is reliably indicated by word order, while in the Latin sentence *Johannes crocodilum interfecit* it is the endings *-s* and *-m* which serve this purpose; however, on hearing an English-based protolanguage utterance such as *John crocodile kill* one cannot know whether to mourn or rejoice without the help of contextual or background knowledge.

One striking fact about hominid evolution is that increase in brain size was not steady. Rather, there was a first burst of brain expansion between 2 and 1.5 million years ago, as *homo habilis* and *homo erectus* came to replace the earlier australopithecines, followed by a second burst within about the last 300,000 years, as *homo sapiens* came to replace *homo erectus* (Aiello 1996a). Various factors, such as diet and group size, have been invoked to explain this. Bickerton's approach to the problem is to ask why *homo erectus*, though capable of quite sophisticated toolmaking, failed to make any significant technological or cultural advance for over a million years. His answer is that *homo erectus* was endowed not with "true" language but only with protolanguage. Those hominids were at least as aware of social relationships as present-day apes are, and could represent thematic structures (who did what to whom) mentally; but they had no reliable linguistic tool for talking about these relationships or expressing these mentally represented structures. Linguistically, they were trapped throughout their lives at the two-word stage of the modern toddler.

Bickerton thus provides an intriguing, though speculative, answer to the question of what held *homo erectus* back for so long. But how did humans ever get beyond protolanguage? Bickerton's answer is that new neural connections in the brain allowed speech to be hooked up to thematic structure. This would have yielded a sudden and dramatic improvement in the reliability and versatility of language, and hence set the stage for the rapid advances of the last quarter of a million years. What is less clear is why the neural hookup should have occurred when it did, rather than earlier or later. Bickerton's scenario also supplies no particular reason why grammar should be more closely associated in the brain with control of the vocal tract than with the organization of vocabulary. But his proposals certainly suggest one way of reconciling the Chomskyan view of modern human language as qualitatively unique with the need to accommodate it somehow in an account of human evolution.

## 6.2 *Actual grammar versus conceivable grammars*

Is the sort of grammar that languages have the only kind that they could conceivably have, or does grammar-as-it-is represent only one of many directions which linguistic evolution might have taken? This is an ambitious question, and there is no guarantee that it can be answered; however, it is the sort of question which only linguists, among the various contributors to language evolution studies, are equipped to tackle.

If one says that the characteristics of grammar-as-it-is are inevitable, one is saying in effect that grammar is as it is for the same sort of reason that the cells in honeycomb are hexagonal rather than (say) square. It would be futile to look for a hexagonal-cell gene in the genetic endowment of bees. This is because, when numerous creatures are trying to build cells in a confined space, each exerting the same outward pressure in all directions, the outcome will



inevitably be a hexagonal pattern for reasons not of biology but of physics and mathematics. That is the kind of possibility that Chomsky has in mind when he says that the ability to learn grammars “may well have arisen as a concomitant of structural properties of the brain that developed for other reasons” (quoted by Pinker 1994: 362). Chomsky has not looked for such structural properties of the brain himself, preferring to concentrate on Universal Grammar itself rather than on what may underlie it. But some researchers are now using computer simulation to explore what happens when a signaling system with certain initial characteristics is set up to be adaptable so as to fit better the needs of the system’s “users” (Batali 1998, Berwick et al. 1998, Steels 1997). If common trends emerge from these experiments, and if these trends correspond to identifiable aspects of grammar and vocabulary, that may indicate that the aspects in question were bound to evolve as they have, irrespective of the fact that it is in the language of humans that they appear rather than in a “language” used by dolphins or Martians. Any firm findings in this line lie in the future, however.

What of aspects of grammar which are not inevitable in this sense? A central issue is whether or not all aspects of grammar are well-engineered responses to selection pressures to which humans are subject. Modern evolutionary theory by no means requires the answer yes. Many characteristics of organisms are mere byproducts of historical accident, and some characteristics are badly engineered for the purposes which they serve. An example is the mammalian eye, in which light has to pass through nerve fibres before it reaches light-sensitive tissue, and the optic nerve causes a blind spot at the point where it passes through the retina (Williams 1966, 1992). (Octopuses’ eyes are more efficient from this point of view.) Natural selection can only tinker with what is genetically available, and perfect outcomes are often beyond its reach. So how much of grammar is well engineered, and how much of it is less than perfect owing to historical constraints?

Pinker and Bloom (1990) and Newmeyer (1991) are inclined to emphasize the positive aspects of grammatical engineering. That is understandable against the background of Chomsky’s emphasis on neutral or even negative aspects. A different tack is taken by Carstairs-McCarthy (1998, 1999), who argues that the grammatical distinction between sentences and noun phrases, despite its familiarity and apparent inevitability, is in fact a piece of mediocre engineering, reflecting the cooption for syntactic purposes of neural mechanisms which evolved originally for the organization of the speech chain into syllables. He suggests that many of the syntactic habits of sentences, verbs, and noun phrases are reflections of the phonological habits of syllables, syllable nuclei (usually vowels), and syllable margins (consonants). This view is consistent with the proximity of Broca’s area to the oral portion of the motor strip, as well as the frequent coincidence of grammatical and phonetic symptoms in Broca’s aphasia. The invocation of imperfections in linguistic engineering as clues to the evolutionary origin of language is quite new, however, and it remains to be seen how fruitful it will be.

## **7 Conclusion**

This tour of recent work on the origins of language has revealed few solid, uncontroversial conclusions. Nevertheless, the field is entering an exciting period. The long freeze in relations between linguists and other language origin researchers is at last beginning to thaw, just when discoveries in archeology, anthropology, primatology, and brain science are all helping to shed new light on the topic from a variety of directions. Will the evolution of language eventually come to be seen by linguistic theorists as not merely a quaint sideline but an essential source of evidence about why Universal Grammar is as it is? My guess is that it will, though the process may take a decade or more. Certainly, the justification for the Paris Linguistic Society's ban no longer exists.