



Why did language develop?

J.F. Stein

University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

Abstract

Language developed for communication, to facilitate learning the use of tools and weapons, to plan hunting and defence, to develop a “theory of mind” and the tools of thought, and to attract and keep a mate. The adaptations required took place over many millions of years. The first important one was left-sided specialisation of the neural apparatus controlling involuntary emotional vocalisations that began more than 200 million years ago. The next was the development in primates of “mirror neurones” in the pre-motor cortex some 45 million years ago. These enabled the imitation and voluntary control of previously involuntary manual gestures and vocalisations. The third important adaptation was the descent of the larynx, 100,000 years ago, which greatly increased the phonological range of vocalisations that could be made. Thus, language did not develop all at once as suggested by Chomsky, but evolved gradually building upon adaptations originally meeting quite different needs.

© 2003 British Association for Paediatric Otorhinolaryngology (BAPO). All rights reserved.

Keywords: Gesture; Mirror neurones; Hemispheric specialisation; Imitation; Chomsky; Vocalisations

To ask why language developed might seem a very odd question; most people would say that the answer is blindingly obvious. Clearly, humans evolved language in order to be able to communicate with each other. Thinking about it long after the event, that seems trivially obvious to us. But we have to ask why more efficient means of communication than already existed in the form of facial expressions, gestures and vocalisations, was so selectively advantageous to humans. Actually language seems to have evolved relatively recently, perhaps only 50,000 years ago. So, despite its obviousness to us, development of language communication was not that overwhelmingly essential. In this paper I want to show that language and then literacy only evolved gradually, long after many other things were in place. Contrary to Chomsky’s suggestion [2], recent opinion is that language was

E-mail address: john.stein@physiol.ox.ac.uk (J.F. Stein).

not the result of a sudden mutation endowing us with a new “neurologically encapsulated linguistic processor”; instead, it evolved gradually from anatomical and physiological adaptations that took place long before. The main reason, I believe, this is important is that it means that we can hope to understand impairments in the development of language and literacy, in terms of the basic biological processes that underlie them. Ontology does to some extent mimic phylogeny, and understanding how basic motor, auditory and visual processes led to the development of language and literacy can help us to elucidate how they go wrong in conditions such as developmental dysphasia and dyslexia.

To understand how language evolved, we need to consider not only the selection pressures that operated to advantage it, but also the genetic variations that occurred by mutation that made it possible. First, therefore, I will provide a very brief overview of ideas on why there were advantages to improving communication by the development of language, and then an equally brief review of the evolution of *Homo Sapiens*. Then, I will describe the evolutionary adaptations that were essential for the development of language before detailing how these may have mediated it.

There are three strains of theory about what kind of selective advantage linguistic communication provided for hominids [1]. The first suggestion is that it enabled the shared use of tools, i.e. the development of technology. In this scenario, the development of language was required to explain how tools should be used and to discuss the best ways of employing them. If these tools happened to be weapons, then they could be used for hunting either animals for food or indeed, enemy Neanderthals. Sad to say, it seems that our ancestors simply annihilated their main competitors.

It has been pointed out that chattering away would not lead to a very successful hunt; but, of course, the conversations would have taken place in order to plan the hunt. The requirement for such planning gives rise to the second main suggestion about the pressure to develop language, namely that language enables you to discuss what is likely to happen by predicting and representing it in abstract terms, i.e. language gives you the tools to think about it.

This representation need not be confined to planning a hunt, but could also be used as the basis of thinking about all manner of things. Perhaps the most important of these would be helping to determine what others are thinking, developing a “theory of mind”. Cynics believe this to be the basis of “Machiavellian intelligence”, working out what people are likely to be thinking, thus being in a better position to deceive them and gain advantage. But, I prefer to emphasise how developing a theory of mind would naturally lead to the invention of the tools of thought: categorisation, quantification, abstraction, causality and logic.

The third kind of pressure that might have favoured the development of language was the requirement to find a mate. This sexual pressure is certainly the main driving force behind bird song. In fact, the left-sided brain structures that control song production by the syrinx actually increase in size during the breeding season; they grow new neurones under the influence of testosterone, and then during the winter, lose these neurones and regress. In addition, because of the long childhood of humans, females have the need to retain their mate for several years to help bring up the children. This is, therefore, the only account that can explain why females are in general slightly better at communication than males.

These three suggestions are not mutually exclusive; probably all three are partly correct. But they are almost impossible to confirm or refute. Hence, they are to a certain extent like Kipling's Just So Stories: how the rhinoceros got his skin might have been because the Parsee put itchy raisins into it when he'd taken it off to swim; but that is very unlikely. Our stories about the selective advantage of communication by speech sound much more plausible, but since unprovable they might be just as totally off beam.

We must now consider the fossil record indicating how humans evolved, incomplete as it is. Mammals first appeared on the earth about 200 million years ago; primates around 65 million; the first Anthropoids about 45 million and Chimpanzees some 12 million years ago. The first Hominoid, *Ankarapithecus*, split from the Chimp line about 9 million years ago, *Australopithecus* appeared around 3 million years ago, *H. erectus* at 2 million, Neanderthals 250 thousand and *Homo Sapiens* 200,000 years ago.

However, the important features of this history from the point of view of language development were that the left side of the brain became specialised for the production of sounds many millions of years ago. Then bi-pedalism developed in the first phase of hominoid evolution, and this freed the arms for more efficient food gathering, tool use and gesturing. Descent of the larynx then made the voicing of phonemes possible to increase their number, and finally, almost in modern time, the development of consistent right-handedness seems to have been important for the invention of writing.

The origins of left-sided cerebral specialisation can actually be discerned from almost the very beginning of life in that the laevo-isomers of biological molecules are always favoured. The strongest asymmetries favouring the left are certainly found in birds, but they are not seen so strongly again until we reach *Homo Sapiens*. It is surely no accident that the two groups that make the most use of acoustic communication site the control system on the left.

However, walking on two legs only commenced with *Homo erectus*, just 2 million years ago, unless you count Kangaroos that evolved in Australia separately from the mammals. The third important adaptation that enabled the development of true speech was the descent of the larynx, but this only occurred about 200,000 years ago. Even though other evidence suggests that true speech did not evolve until much later, perhaps only 35,000 years ago, descent of the larynx and the wider variety of sounds that this enabled must have been sufficiently advantageous to outweigh the danger of inhaling food.

The order in which these adaptations appeared is important because it sketches out the basis of how the capacity for speech may have developed innately. For a long time there have been arguments about whether language is innate or learnt. In 7000 BC, the Egyptian pharaoh Tsammtemichus was said to have had a child brought up in isolation in a cage to find out whether he would speak without any teaching. His first word was the Phrygian word for bread; which he'd clearly learnt from his guards, so the experiment suggested that language was learnt. In Agra, in the 17th Century, AD Mogul Akbar Khan ensured strict silence in his brutal version of the experiment by employing dumb nurses to rear 12 children in isolation. He found to his surprise that none of the children learned to speak at all, again implying that language had to be learnt by example. In the 1960s, Genie was found in Los Angeles completely unable to talk because she had been protected by her over religious parents from almost all sensory input that might have contaminated her with evil, by being locked up in a dark cupboard. These and a host of other evidence have

shown clearly that language has to be learnt. The English learn English from their parents; the French learn French.

On the other hand, Johann, who had been abandoned as a baby in Burundi, was brought up by chimpanzees and when found at the age of 5 was using chimpanzee vocalisations and gestures to communicate. This shows that our evolutionary past equips us to learn languages but does not provide language itself. Nor does it specify the language we learn. In Stephen Pinker's words, our genome gives us an "instinct" to communicate, but we have to learn the means to express it.

What we now have to do is to sketch out the ways in which this came about [3]. Our starting point will be the vocalisations that almost all mammals produce. As in birds, these are controlled by a network of neurones that shows a propensity to favour the left-hand side. Why it is the left side that is chosen we do not know, and even why one side is favoured is not entirely clear. Most features of animals are bilaterally symmetrical and perfectly satisfactorily controlled from both sides of a bilaterally symmetrical brain because there are good cross communications between the two sides.

Nevertheless, in about 95% of humans (including 70% of left-handers), parts of the left cerebral hemisphere are specialised for mediating the perception and production of speech. Probably one side is chosen because if the two hemispheres both try to control the same structure, they tend to compete with each other; hence, the majority of the cross connections between the hemispheres have been found to be inhibitory in order to prevent the two sides trying to do the same thing. Placing the control of vocalisations predominantly in the same hemisphere thus simplifies the control problem. Another contributory factor may be that the length of axons joining sensory and motor language areas within one hemisphere would be slightly shorter than requiring connections between the two hemispheres, thus cutting down on delays in a system that requires millisecond accuracy.

In lower mammals, including monkeys, vocalisations are controlled by a medially placed system of neurones that involves the cingulate cortex, basal ganglia and hypothalamus. The composition of this system suggests strongly that it is primarily concerned with the expression of emotions, and it is now clear that in primates other than man all vocalisations are automatic, driven by the emotions. Thus, all attempts to teach primates to actually talk have failed; it is impossible to harness primate vocalisations for other kinds of communication because they are not under the animals' voluntary control. Chimpanzees are actually intelligent enough to be taught to communicate to some extent using their extensive repertoire of voluntary gestures, but never by vocalisations.

It is highly significant that this emotionally controlled medial system does not involve the monkey homologue of Broca's speech area in the left lateral frontal cortex. Indeed, lesions in this area do not affect the ability of monkeys to make their vocalisations at all. The development of the human motor speech area has followed a different course. This has been powerfully illuminated by the discovery of "mirror" neurones [5]. These unequivocally place the development of human language in the province of gesture and facial expression.

Mirror neurones are found in the ventrolateral frontal lobe just in front of the face and arm representation in the primary motor cortex. Their important characteristic is that they fire not only when a monkey reaches out to grasp an object, but also when the monkey observes somebody else doing the same thing; however, not when the same goal is

achieved in a different way, for instance using a pair of pincers. In humans, likewise, this area appears to be activated both when the subject reaches to grasp something, but also when he imagines doing it and also when he sees somebody else doing it. Thus, mirror neurones could underlie how we learn to produce speech by enabling us to imitate our parents' speech. In addition, they offer unexpected support for Lieberman's motor theory of speech perception [4]. Mirror neurones would enable us to interpret speech because the very same cells would be activated by observing speech as those that we would employ to make the same speech sounds.

In addition, Corballis and others have argued very convincingly that these mirror neurones show that speech evolved from gesture and not from vocalisations [3]. The argument runs as follows: by representing a particular gesture, mirror neurones enable other people to imitate it and thus to communicate by means of these gestures. Lieberman, Studdert-Kennedy and their colleagues at the Haskins laboratory of speech science have long argued that the basic elements of speech are not, as generally assumed, consonants and vowels, but rather the vocal gestures that generate them, namely the movements of the lips, tongue and larynx [4]. Thus, mirror neurones in Broca's area could come to represent lip, tongue and laryngeal gestures, and these could generate the phonetic elements of speech. In fact, most of us still gesture with our hands when speaking and when we gesture our vocal production is synchronised with our hand gestures. These voluntary movements of the vocal tract are mediated by this lateral system, and they have been superimposed on the older emotional system for automatic vocalisations mediated by more medial brain structures. The latter still supply the basic intonation and prosody for sentences.

This account of the development of speech from gesture has received further support from the recent discovery that the gestures of sign language used by the deaf are controlled by the same left hemisphere centres, particularly Broca's area, that speech occupies in those who can hear. One can even see in the way in which the order of ideas and syntax is expressed in the trajectory of a sign language gesture, how language syntax and grammar may also have been founded in the way that the structure of a signed sentence is determined by the evolution of a gesture from shoulder to fingers.

The final step in this history is to consider the invention of writing. This is truly a cultural invention, and probably not at all enshrined in our genome because it was only invented about 5000 years ago and was not common until the last century. Being able to read and write carried no particular selective advantage; and, it is therefore most unlikely that we would ever find a gene or genes for reading, contrary to what is sometimes claimed. But, like speech before it, writing depends on prior adaptations, in particular the development of articulatory gestures controlled from Broca's area. In addition, its invention probably depended on the development of right-handedness. Despite the choice in most animals of left-brain structures for the control of both automatic emotional vocalisations and voluntary speech, no lower animal shows such strong right-handedness as humans do. Many animals choose to use either left or right hands for particular tasks, but not even chimpanzees choose the right so systematically.

Probably the main reason why right-handedness was so important for the invention of writing is that hieroglyphs and letters are very impoverished visual signals, and it matters greatly whether they are pointing to the left or to the right. It is easiest for us to agree which way round they should be if we all write from the same side. Ambidextrous

animals and children therefore have extreme difficulty knowing which way round bs and ds should go.

This is not to say that right-handedness only evolved 5000 years ago. Many of the Stone Age hand axes from 100,000 years ago show signs of having been shaped for right-handers. Likewise the beautiful cave paintings of Lascaux and Altamira flowered at about the time that speech evolved; but most of the people depicted seem to have been right-handed. Like speech, therefore, writing has clearly piggybacked on an adaptation that occurred a lot earlier for a different purpose, perhaps gesture.

Reading and writing are much more difficult to learn than speaking is because the written word does not map so easily onto articulatory gestures as speech does. As Lieberman and Studdert-Kennedy point out, the phonemes that are represented by letters are artificial subdivisions of the articulatory gestures that generate them, and these subdivisions have to be taught and learnt. Hence, a very large proportion (some 10–20%) of humans never master this art properly; it is the most difficult thing that most of us ever have to learn.

Let us now return to Chomsky [2]. We can now see that his idea that the human acquisition of language and literacy resulted from a single mutation that endowed us with a linguistic processor, all in one jump, is clearly inconsistent with recent discoveries. Speech and language evolved gradually, coat tailing on a series of adaptations that evolved for completely different purposes and occurred millions of years earlier.

Nevertheless, Chomsky was not entirely wrong. I've always been a great admirer of him, and his insights concerning phonology, deep grammar and syntax come out of our current concepts of the development of language rather well. The pre-eminence of phonology follows from the descent of the larynx and the specialisation of Broca's area for the voluntary control of articulatory gestures. Syntax and grammar can be viewed as a direct consequence of the evolution of sentences out of gestures that can involve the whole body from axial back muscles to distal finger-movers. In rather the same way that identical Chinese logographs can represent totally different sounding words in Chinese and Japanese, so the same deep structure, the same flow of ideas conveyed by a particular gesture, could be represented by different strings of articulatory gestures producing the thousands of different languages that have developed throughout the world.

What makes this whole enterprise more than just curiosity is that it provides new insights into what can go wrong with language. Because ontogeny repeats phylogeny to some extent, elucidating how language gradually evolved from archaic gestures means that potentially we now have a new approach to understanding developmental disorders of language. Counter-intuitively, both speech production and comprehension seem to depend on our mirror neurones being able to represent articulatory gestures. Hence, although comprehension clearly makes greater demands on the auditory system, and speech production makes greater demands on Broca's area and the motor vocalisation system, Broca's area is engaged in both, and this is what modern imaging methods have shown clearly. As expected therefore, children with developmental dysphasia are significantly worse at deciphering articulatory gestures, as in lipreading. Also, whereas good speakers' hearing of a phoneme is greatly altered if the speaker's lips appear to be generating a different one, the McGurk effect (in developmental dysphasia, the mishearing of the

phonemes) is much less apparent, because in them the mirror system is not working as well as it ought to be.

References

- [1] J.L. Bradshaw, *Human Evolution*, Psychology Press, Hove, 1997.
- [2] N. Chomsky, *Reflections on Language*, Pantheon, New York, 1975.
- [3] M.C. Corballis, *From mouth to hand*, *Behav. Brain Sci.* (in press).
- [4] A. Liberman, D. Shankweiler, M. Studdert-Kennedy, Perception of the speech code, *Psych. Rev.* 74 (1967) 431–461.
- [5] G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, Premotor cortex and the recognition of motor actions, *Cogn. Brain Res.* 3 (1996) 131–141.