

# **Preconditions for the evolution of protolanguages**

**Merlin Donald  
Department of Psychology  
Queen's University  
Kingston, Ontario  
K7L 3N6, Canada**

**E-mail: [donaldm@psyc.queensu.ca](mailto:donaldm@psyc.queensu.ca)**

***In: The Descent of Mind: Psychological perspectives on hominid evolution.***

# The Descent of Mind

Psychological Perspectives on Hominid Evolution

---

Edited by

**Michael C. Corballis**

*Department of Psychology,  
University of Auckland*

and

**Stephen E. G. Lea**

*Department of Psychology,  
University of Exeter*

**OXFORD**  
UNIVERSITY PRESS

---

# Preconditions for the evolution of protolanguages

Merlin Donald

---

## Introduction

Most theorists in the field of language evolution concede that the earliest forms of language must have emerged from adaptive elaborations on ape gesticulation and vocalization, and that these skills were gradually transformed by evolutionary pressures into the more flexible, and ultimately much more powerful, means of communication that we know as human language. There is such widespread agreement on this basic idea, from Darwin (1871) to Lieberman (1984), Bickerton (1990), Pinker and Bloom (1990), Corballis (1991), Greenfield (1991), Pinker (1994), and many others, including myself (Donald 1991), that one must conclude that, within the bounds of a neo-Darwinian theoretical framework, no viable alternative to this position has yet been proposed. The only exception of any influence is an extreme anti-Darwinian position once expressed by Chomsky (1965), that human language is based on a principle entirely different from that of animal communication. In essence, Chomsky's position holds that there are no precedents for language, and that language must have emerged by means of some discontinuous process that cannot be encompassed within the current neo-Darwinian framework. Unfortunately, this position does not generate dialogue, and is more an *ex cathedra* proclamation than a theory.

I will assume for the purposes of this article that the evolution of language, like all other aspects of life, must have been subject to the laws of natural selection, and that its emergence must therefore have been a continuous process. Given this assumption, there is still considerable room for disagreement on the details of the evolutionary transition from ape to human, and on the form that a credible theory of language evolution should take. There are early-versus-late language models; 'punctuated' models, with discrete stages of evolution, versus gradualistic models; vocalization-first versus language-first models; 'ontogeny-recapitulates-phylogeny' versus glottogenetic models; and encephalization versus lateralization models, to list just a few. The diversity of language-origin proposals reflects the fact that we have virtually no direct evidence on precisely what the cognitive skills of early hominids might have been. However, this diversity might be narrowed down if we consider the cognitive constraints that must be imposed on theories of language evolution.

The term 'protolanguage' has been used as a label for the earliest form that 'true' language must have taken. A protolanguage is thus a prototype, an approximation of the finished product that contains some of the essential features of language, and that could later have evolved into full-fledged language. Based upon his work on pidgins and creoles, Bickerton (1990) has proposed a specific definition of this phenomenon. He argues that protolanguage must resemble pidgins, which are very much like the language of two-year-old children, consisting of one- and two-word utterances combined with pointing, gesturing, and prosodic voice modulation. Pidgins are also grammarless. Bickerton proposes that protolanguage must have had all these characteristics. He locates protolanguage far back in hominid evolution, with the earliest members of species *Homo*, who lived in Africa over two million years ago. In line with his theory, he argues that grammars came later in hominid evolution, with *Homo sapiens*. He builds a credible case for this scenario, pointing out how a rudimentary vocabulary would have had immediate adaptive advantages for pre-sapient *Homo* (i.e. *Homo* before *Homo sapiens*), in tool-making, hunting, and social coordination.

Bickerton's approach thus places language early; it has an element of gradualism in it, although it acknowledges a period of accelerated change associated with the speciation of sapient humans; and it falls into the class of 'ontogeny-recapitulates-phylogeny' theories alongside those of Parker and Gibson (1979), Greenfield (1991), and others. But underneath these surface features there lies an assumption, one might say an axiom, that drives his particular theory, as well as a number of similar ideas about language origins.

The crucial axiom supporting Bickerton's argument is that *language is fundamental to all distinctly human thought and consciousness*. The root syllogism leading to his theory can be reconstructed thus:

- (1) archaeological evidence shows that the culture of pre-sapient *Homo* was a major advance over what preceded it, manifesting the germ of human inventiveness, craft, and social coordination;
- (2) this, combined with a large increase in the encephalization ratio of pre-sapient *Homo* implies a significant cognitive advance in the direction of symbolic thought;
- (3) and symbolic thought is inherently dependent upon language capacity (Bickerton 1993).

Therefore, some primitive form of language skill must have been present in *Homo* before the emergence of *H. sapiens*.

A constraint acknowledged by Bickerton is that we must withhold some aspect of cognitive evolution from pre-sapient *Homo* to explain the tremendously accelerated recent cognitive advance of *H. sapiens*. For Bickerton, the extra advantage that sapient humans enjoy is the capstone of the Chomskyan conception of the linguistic world—grammar. Thus, Bickerton not only explains the achievements of pre-sapient *Homo*; he also preserves the most important component of Chomsky's treasured 'language acquisition device' for sapient humans.

Bickerton's approach is compatible with Daniel Dennett's belief that thought and consciousness are a direct product of language capacity. Dennett (1992) addressed this specific point, arguing that some form of language would have been necessary for

pre-sapient hominid skills like fire-tending and complex tool-manufacture. In general, one might say that something like Bickerton's version of evolution is favoured by those who believe that human intellectual advantage is exclusively tied to language ability. This would probably include most classical A/I (Artificial Intelligence) theorists, many philosophers, and most linguists, who, not surprisingly, tend to be very language-centred in their bias. In this world-view, to explain the great cognitive leap that was obviously taken by pre-sapient *Homo*, one must push language back into that era.

I disagree with this world-view. The roots of my disagreement run very deep, and can be traced to the axiomatic level. Returning to the implicit syllogism I attributed to Bickerton, I would agree with its first two propositions, but not with the third, adding a supplementary caveat. The new (altered) argument would take the following form:

- (1) archaeological evidence shows that the culture of pre-sapient *Homo* was a major advance over what preceded it;
- (2) this, combined with a large increase in the encephalization ratio of pre-sapient *Homo*, implies a significant cognitive advance;
- (3) such an advance would not have been possible without a more powerful form of representation; and then the caveat:
- (4) *we must not assume that this early advance was in itself linguistic*, or that it conformed to conventional definitions of symbolic thought.

We should favour a theory that pays attention to the cognitive fundamentals, and does not make the leap to language until it is absolutely necessary. It also follows that we should explore statements (3) and (4) in some detail before trying to draw any conclusions.

### **The primacy of motor evolution**

Could protolanguage have sprung 'fully-armed' from the cognitive armamentarium of primates? My answer to this question is a resounding 'no.' There are important fundamentals missing from the primate mind, without which protolanguages could not emerge; I shall call these the 'cognitive preconditions' of protolanguage. The nature of these preconditions becomes clear when we consider the collective, social nature of language. There are published reports of intentional gesticulation in individual chimpanzees and isolated 'pedagogical' interactions between chimpanzees (see, for example, Boesch and Boesch-Ackerman 1991), but the acid test for a species-wide representational capacity is its *general and spontaneous adoption in culture*, as shown by the presence of significant expressive variability across generations and between different social groups of the same species. This does not happen in modern apes, or in any other documented mammalian species (the variability found in bird-song forms an interesting parallel, but has no direct bearing on primate evolution). An adequate theory of language origin must therefore confront the problem of symbolic invention in groups. To cross the gap between symbolic and non-symbolic representation, hominids needed a capacity to invent expressive conventions—gestures and symbols—'in the wild.'

This focuses the question somewhat—what type of cognitive change would enable a group of primates to invent highly variable and culturally idiosyncratic forms of gesture and sound, and gradually develop distinctive cultures? Again, we must not assume that language was essential to achieve this end. A more fundamental requirement would appear to be a generalized capacity for deliberately refining action. All gestures and intentional vocalizations are ultimately actions of the musculature, and to generate greater varieties of gestures and sounds, primate motor behaviour must somehow have become much more plastic, less stereotyped, and subject to deliberate rehearsal. In other words, a breakthrough in hominid motor evolution must have preceded language evolution.

This revolution in the cognitive control of action required a break with the traditional scenario that evolutionary biology has developed to explain the evolution of animal behaviour. Throughout most of mammalian evolution the range of behaviours available to a species has been largely fixed in the genes, and closely attuned to its ecological niche; this is sometimes called a 'specialist' strategy (cf. Gamble 1993). While this mechanism enables tremendous diversity across species, and some specialization of roles within species, it forces virtual stereotyping on the behaviour of any single species. This is generally true of higher mammals; if you have seen one social group of lions or elephants, basically you have seen them all. The genetic concrete in which their behaviour is cast does not allow animal species to escape from the stereotyped shape of their own motor repertoire; their motor options are confined within a narrow 'morphological envelope'.

It is clear that human languages have broken out of this morphological straitjacket, and are constructed upon the virtually infinite variability of range and arbitrariness of human communicative behaviour. Where would such flexibility have begun? This question leads us logically back to one of the fundamentals of human motor skills, or what is sometimes called procedural learning. To vary or refine an action—any action, not only speech—one must carry out a sequence of basic cognitive operations. These are, traditionally: rehearse the action, observe its consequences, remember these, and then alter the form of the original act, varying one or more of the parameters dictated by the memory for the consequences of the previous action, or by an idealized image of the outcome. We might call an extended cognitive sequence of this sort—whose inherent complexity should not be underestimated—a 'rehearsal loop'. This is the familiar process of conscious review, whereby humans redirect their attention away from the world, and toward their own actions. The focus of attention thus becomes, not the reward or punishment that follows an act, or its social consequences, but the form of the act itself.

Apes seem to lack this option. They might repeat actions as a function of fairly obvious reinforcement contingencies, or socially-facilitated imitation, but they do not initiate and rehearse actions entirely on their own, for the sole purpose of refining their own movement-patterns. An important case in human prehistory is the skill-sequence known as throwing. Various species of primates throw projectiles, but they do not practise throwing projectiles, thus refining their accuracy and power, and creating variants on the execution of this skill. In contrast, even young children routinely practise and refine skills for endless hours, for example experimenting with various styles of throwing. This is a complex process; it demands that the skilled learner

should be able to effortlessly change parameters such as starting position, grip, angle of arm and shoulder, release, velocity, acceleration, and so on, while evaluating the outcome of the modified action-pattern. Calvin (1983) has emphasized that humans need improved ballistic motor skills to become good at throwing, but surely the most critical element is a capacity for deliberately reviewing self-actions, so as to experiment with them. Systematic and repetitive experimentation with action is evident fairly early in human development, especially in infant babbling, including manual sign-babbling. It would be no exaggeration to say that this capacity is uniquely human, and forms the background for the whole of human culture, including language.

Those who believe in the primacy of language and symbols seem to think precisely the opposite, that this remarkable human control over motor behaviour came only after the evolution of language. Bickerton (1993), in a comment on my theory, proposed that, as humans developed better thinking skills (due presumably to protolanguage), they also developed better motor skills in general. In his 'language-first' version of evolution a child that is bouncing a ball off a wall and catching it would be able to do this only because everything in the set piece—the ball, the wall, the acts of throwing and catching—can be captured by language. Because the human child can think about the act linguistically, it is therefore able to experiment with it, and refine it through practice. Following the same logic, as Daniel Dennett (1992) has suggested, early hominids could have fashioned a complex tool only if they could 'talk' to themselves, in some primitive (proto)language, about the actions involved. Hence their conclusion—language in some form comes first, skill second.

But language-first theories run into many difficulties. Apes, and many other species, have thinking and symbol-using skills that are quite complex, without the use of language. For instance, the bonobo Kanzi recently learned how to make and use a primitive version of Oldowan tools (Toth *et al.* 1993). Such demonstrations contradict the notion that all higher cognitive skills depend on language. Second, a large experimental literature has shown that non-verbal representations and procedural learning are orthogonal to the presence of language. This widely-replicated finding has led to the dual-coding hypothesis, the notion that humans have two broadly different intellectual strategies available for representation and thinking (see Paivio 1991 for a review). Moreover, in evolutionary terms, language-first theories ask for a virtual saltation from ape cognition to language—which in the worst scenario becomes a kind of Creationism. Quite simply, it puts the cart before the horse. Where would the ability to think in primitive languages have come from in the first place? As every schoolchild knows, languages are made up of lexicons and grammars, and the sounds and grammars of specific language families are very different from one another. How could archaic hominids have broken out of the stereotypy imposed by the standard primate motor repertoire, and evolved the rich variability of action-patterning that would set the stage for the emergence of language, and even of protolanguage? The answer to this question is clear—the *ability to self-programme motor skills had to come first*. Lexicons could not have come into existence—that is, they could never have evolved 'in the wild'—unless a great advance in motor skill had already taken place. A capacity for the purposive rehearsal and refinement of skill had to come into existence first, and thus it had to come without the pre-existence of

language, with its own evolutionary *raison d'être*, and as I have argued (Donald 1991, 1993a, b, c, 1995, 1996, 1997, 1998) its own representational principle.

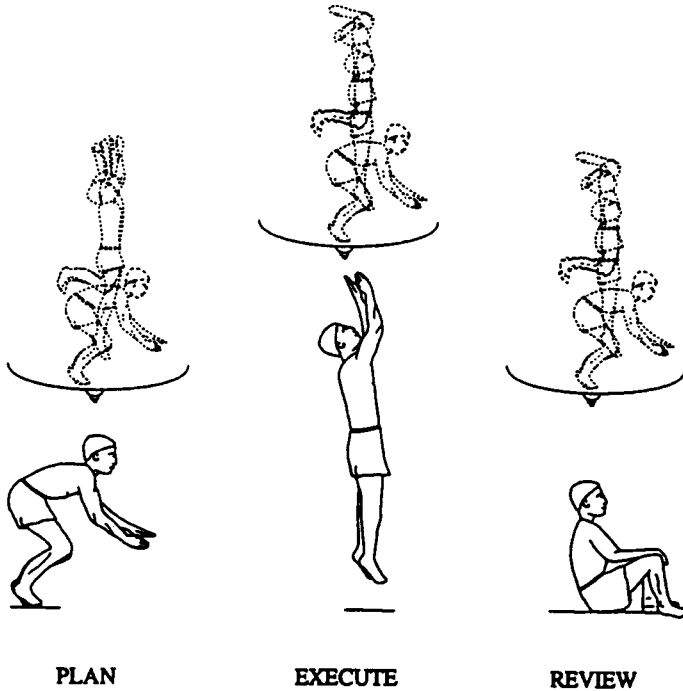
The preconditions of protolanguage were even more demanding than that. The key to purposive rehearsal and morphological variability is voluntary recall: Hominids had to gain *access to the contents of their own memories*. You cannot rehearse what you cannot recall. If an animal depends entirely upon environmental triggers to remember when and what to rehearse, skill-development becomes extremely difficult, since the animal cannot self-trigger the memories supporting the skill, and effectively hangs in suspended animation until the environment provides the cues needed for retrieval of a given response-pattern. Trainers of apes have to cope with precisely this limitation; for instance, it often takes thousands of trials to establish a reliable signing response in a chimpanzee (Greenfield and Savage-Rumbaugh 1990). In general, primates and other mammals have great difficulty in voluntarily self-triggering the contents of their own memory banks, and seem dependent upon conditioning by the environment. Even the enculturated apes of Savage-Rumbaugh *et al.* (1993) rely heavily on the human cultural environment provided by their caretakers for memory cues. Humans, in contrast, are able to 'think about' or 'imagine' things that are totally unrelated to the immediate environment, seeking out specific memory items, preferably without retrieving unwanted elements, and without relying on the environment to come up with the relevant cues to help find the target item.

To be able to do this—to focus selectively on one's own action-patterns, there has to be an *implementable* representation of action in the brain. Agents of actions must be able to trigger action sequences out of their own imaginations. In modern humans we recognize this highly-developed capacity for kinematic imagination; coaches of professional athletes and teachers of musicians and actors rely heavily on this skill. Another key to self-programmed action is that *imagery generated in kinematic imagination may be edited before being implemented in action* (Fig. 8.1). Without this capacity, there could be no refinement of human movement, no increase in its variation within the species. Without self-cued retrieval from memory, under the control of kinematic imagination, humans would remain locked into a reactive strategy, driven by the vicissitudes of the environment, like most mammals. The latter point is important—non-verbal representation is rooted in the human creative imagination, and operates by *analogue* principles. Mimetic meaning is driven by perceptual, rather than linguistic rules. Nevertheless, it is representational, even though, in the simplest case, that of skilled rehearsal, mimetic representations might serve only to represent themselves. In more complex social situations, however, mimetic communications can become highly intricate and complex, as they are in modern human society, without ever invoking a linguistic mode of representation.

### Domain-general models of models—mimesis

These advances in the refinement of voluntary action—kinematic imagination, and a consciously-controlled rehearsal loop—might have taken the form of highly specialized adaptations in humans. For instance, it is conceivable that humans might have evolved a greater degree of specialized manual or vocal control, with the latter





**Fig. 8.1.** The conscious review process. Actors can consciously review their own kinematic imagery before, during, and after execution of an act. This 'rehearsal loop' enables an actor to modify future acts until match is achieved to an idealized template of the act. The latter may be modified either in action or in imagination. This capacity suggests the existence of a consciously-accessible modelling process in the human brain, in which the actor's own movements can be integrated with an event-perceptual model of the external world.

following the same evolutionary route as bird song. But this would have led nowhere in terms of general intellectual progress. Judging from the capacities of modern humans, the advancement of motor skill in hominids was a more domain-general adaptation; that is, it applied to the whole primate voluntary motor repertoire.

There is an opposite viewpoint on this issue, epitomized by the work of Tooby and Cosmides (1989), and recently applied to hominid prehistory by Mithen (1996). These thinkers tend to treat language, and indeed all primate cognition, as if it were divided into cleanly-dissociable modules that evolved along independent trajectories. Thus we are supposed to have distinct social, spatial, and communicative modules, within which there may be even more specific modules, for a variety of specialized functions. It is significant that these theorists rarely cite cognitive or neuroscience research, yet do not hesitate to speculate at length about the organization of both the brain and cognition. Perhaps this is understandable, because the experimental literature contradicts their views. It contains many demonstrations that domain-general capabilities are, and always have been, an essential feature of the vertebrate brain. Tooby and Cosmides seem to have confused the term 'domain-general' with what they conceive as 'general-purpose' capabilities. While it is true that evolution could not, even in

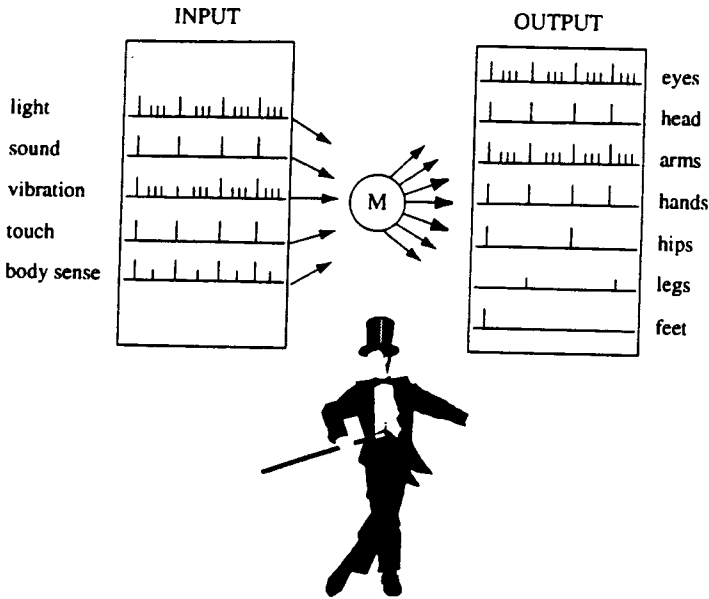
theory, generate general-purpose capabilities, the term 'domain-general' applies to cognitive architecture, not to the process of adaptation. Just as modular ones, domain-general capacities can be adaptations of a specialized nature. Human intentional communication cuts across all major sensory and motor modalities, and by definition must be considered domain-general.

Neurophysiologists sometimes refer to a domain-general control system as a supra-modal system. A supra-modal system is, literally, 'above modality;' that is, it can take its inputs from various sources, and implement its output in any set of voluntary muscles. Humans have excellent supra-modal capacities; thus, I might read the letter 'a' with my eyes, or my fingers, or for that matter, my nose; and I can write the letter 'a' with my finger, my wrist, my elbow, my lips, my head, or even my legs or trunk. This means that my abstract perceptual template for the letter 'a' can be mapped on to virtually any subset of my voluntary musculature. To a degree, supra-modal motor control exists in other mammals, but only for certain fixed classes of action, such as running, eating, or mating. Other mammals do not seem able to build a second-order 'model of models,' in which their own acts become part of the model, and thus subject to systematic rehearsal and refinement. Thus, baboons throw projectiles in a fight, but there is no evidence that they can improve their skill at this by deliberate rehearsal. To achieve this, they would have to shift their attention away from the external world, and redirect it towards their own bodies and actions, and they seem incapable of this. In contrast, even very young human children can spontaneously rehearse such actions, including apparently non-utilitarian actions such as standing on one's foot, making faces, or skipping stones.

The most compelling manifestation of this uniquely human domain-general motor skill is found in motor rhythms. Motor rhythms are transferable to any skeletomuscular system in the body, singly or in combination, and this illustrates the abstractness of the schemata that drive them. A rhythmic source—usually auditory, proprioceptive, or visual—might be tracked vocally, or with the fingers, the feet, or the whole body (Fig. 8.2). The controlling construct is an abstract temporal model that can be translated into various concrete kinematic images: For instance, in a jazz drummer's improvisations; in dance or marching; in coordinated group song; in many children's games; or in gymnastics.

The process that generates these action-patterns relies on a principle of perceptual resemblance; accordingly I have labelled this skill 'mimesis' or 'mimetic skill'. Mimesis is essentially metaphoric, or holistic. A mimetic action—for instance, a rhythm—can be grasped effortlessly in many variations and approximations, and the pattern cannot easily be reduced to a computational principle; that is, it cannot always be reduced to discrete or digital elements that are combined according to rules. Rather, it is more like the visual recognition of faces—the 'Gestalt,' or overall pattern, is primary.

This holistic feature distinguishes my notion of mimetic skill from Corballis's (1989, 1991) idea of 'generative praxis,' which operates according to computational, rather than holistic principles. Corballis's praxic adaptation was supposed to segment or parse action; that is, break it into irreducible fundamental components, and recombine these components at will. Thus, a tool-making sequence might be reduced to a few basic action-elements that are essentially symbolic in nature, and can be combined and recombined in various ways. These elements in turn can be parsed into



**Fig. 8.2.** The integration of mimetic expression with external stimuli. In rhythmic actions such as dance the actor maps from any or all of several input modalities to any or all of several output, or motor, modalities. Thus the whole body becomes an expressive device. This is, by definition, a domain-general capacity. These motor-mappings can be switched, altered, or recombined by the actor at very short notice. The acquisition or modification of such mappings is usually achieved during conscious rehearsal, using active kinematic imagery.

recombinable components. Mimetic action does not operate in this way. It is extremely difficult to reduce vocomimetic expressions, body-language, metaphoric gesture, or a complex re-enactive sequence, to discrete elements. Undoubtedly, the inspiration for Corballis's idea, and for Kosslyn's (1988) closely-related conceptualization of visual scene-analysis, is the modern computational habit of parsing and labelling events using explicit symbols. But just because programmers can find no alternative to using language-like symbols to parse perceptual events, it does not follow that the mammalian brain does the same. The denotative uses of language that support programming are extremely recent innovations, and a very unlikely source of their own invention.

Human mimetic capacity can model complex events that extend over fairly long time-scales, to include the sequencing of more complex patterns of action in context. Thus, if hominids could comprehend and remember a complex event, such as the killing of an animal or the manufacture of a tool, they should have been capable of re-enacting such events, individually or in groups, once mimetic capacity was established. Such re-enactments would have served the elementary pedagogical function of transmitting the sequence: Thus hunting and tool-making techniques could be extended and transmitted throughout the group. To some extent, this sort of imitative behaviour is present in apes, and they obviously understand many complex sequences when they are demonstrated by humans; but evidently they lack the critical capacities

needed to ignite the spontaneous cultural innovations that mark human culture. Representationally-driven cultures are impossible without a public arena of modifiable action. This emphasizes once again the primacy of output—that is, production systems and ultimately motor control—in human evolution.

### **The cultural impact of mimetic skill distributed in social groups**

An improvement in primate motor skill of this magnitude would inevitably have resulted in changes to hominid patterns of social expression, but it would not necessarily have led to anything that could qualify as protolanguage. The entire existing repertoire of primate expressive behaviour would have become raw material for this new motor-modelling mechanism. By 'parachuting' a domain-general device of this power on top of the primate motor hierarchy, previously stereotyped emotional expressions would have become rehearsable, refinable, and employable in intentional communication. This would have allowed a dramatic increase in the variability of facial, vocal, and whole-body expressions, as well as in the range of potential interactive scenarios between pairs of individuals, or within larger groups of hominids. This is precisely what we can see in modern humans. Importantly, because a supra-modal mimetic capacity would have extended to the existing vocal repertoire, it would have increased selection pressure for the early improvement of mimetic vocalization, a skill whose modern residue in speech is known as prosody.

Given a mechanism for intentional rehearsal and refinement, instrumental skills would also have moved to another plane of complexity through sharing and cultural diffusion. Sophisticated tool-making is in many ways the most solidly-established achievement of archaic *Homo*, but it is important to realize that the manufacture of a new kind of tool implies a perceived need for that tool, and corresponding advances in both tool use and pedagogy. Mimetic skill would have enabled widespread diffusion of new applications, and supported the underlying praxic innovations that led to new applications.

In addition to tool-making and emotional expression, motor mimesis would inevitably have allowed some degree of quasi-symbolic communication, in the form of a very simple shared semantic environment. The 'meaning' of mimed versions of perceptual events is transparent to anyone possessing the same event-perception capabilities as the actor; thus mimetic representations can be shared, and constitute a cognitive mechanism for creating unique sets of socially distributed representations. The expressive and social ramifications of mimetic capacity thus follow with the same inevitability as improved constructive skill. As the whole body becomes a potential tool for expression, a variety of new possibilities enter the social arena: complex games, extended competition, pedagogy through directed imitation (with a concomitant differentiation of social roles), and public action-metaphor, such as intentional group displays of aggression, solidarity, joy, fear, and sorrow. These would have perhaps constituted the first social 'customs,' and the basis of the first truly distinctive hominid cultures. This kind of mimetically-transmitted custom still forms the background social 'theatre' that supports and structures group behaviour in modern humans.

The emergence of mimetic skill would also have amplified the existing range of differences between individuals (and groups) in realms such as social manipulation, fighting and physical dominance in general, tool-making, tool use, group bonding and loyalty, pedagogical skill, mating behaviour and emotional control, and even in sexual behaviour, which would have become much more complex. This would have complicated social life, placing increased memory demands on individuals; but these communication tools would also have created a much-increased capacity for social coordination, which was probably necessary for a culture capable of moving a seasonal base camp or pursuing a long hunt.

It is important to consider the question of the durability of a hominid society equipped with mimetic skill—adaptations would not endure if they did not result in a stable long-term survival strategy for a species. Mimesis would have provided obvious benefits, allowing hominids to expand their territory, extend their potential sources of food, and respond more effectively as a group to dangers and threats. But it might also have introduced some destabilizing elements, especially by amplifying both the opportunities for competition, and the potential social rewards of competitive success.

### **Mimesis as a pre-adaptation for protolanguage**

All these changes could have occurred without any form of what might properly be called 'linguistic' modelling of reality. In my version of events, the metaphorical, holistic pattern of thought that marks mimetic imagination would not have had the capacity to parse the visual display and label its component parts; it would have had no need for lexicons. Nevertheless, mimetic skill provided an essential pre-adaptation for speech and language, for at least one very obvious reason: the *primacy of mimetic skill in lexical morphophonology*.

Mimetic skill was necessary for the later evolution of language, because the evolution of mimetic skill made it possible to tamper with the morphology of action. But this development was only one of the two major preconditions for lexical invention, and thus for protolanguage. We should perhaps reflect on the nature of the lexicon for a moment. I will use Levelt's (1989) schematization of the lexicon as a basis for discussion. A lexical entry has to have two major components: the 'lemma', which contains the semantic and grammatical aspects of each word, and its morphophonology. In principle, the morphophonological components of a lexicon could have been constructed along purely mimetic lines. But the lemma would not have been provided by a mimetic motor adaptation, and without this side of the lexical entry, there can be no lexicon in the normal meaning of the word. There might be elements of 'speech acts' that could be supported mimetically; and there might be gesture, including vocal gesticulation; but this does not constitute a basis for constructing a true lexicon.

Mimesis was therefore necessary, but not sufficient; however, this does not diminish its importance in the later evolution of language. To invent a primitive lexicon, even in the halting, grammarless manner of a two-year-old, one must be capable of a considerable degree of mimetic invention and refinement. To be more precise, one must be able to produce *retrievable action-schemata*. Before any form of true oral

language could appear, hominids had to be able to create large numbers of retrievable (or what I call 'autocuable') motor schemata. These would have provided the potential for developing a repertoire of lexical entries, the *sine qua non* of any protolanguage system.

To reiterate my major point, the fundamental principle of self-triggered voluntary retrieval of motor representations had to be established in the brain before the highly complex motor acts of speech would have been possible. Phonetic skill has been called 'articulatory gesture' by various investigators (Brownman and Goldstein 1989); the whole higher apparatus of speech depends on the basically mimetic ability of individuals to create rehearsable and retrievable vocal acts, usually in close connection with other mimetic acts. The same principle applies to sign language. Indeed, the iconic grammars of sign language have an even closer link to mimesis than the grammars of speech. This linkage has led Armstrong *et al.* (1995) to propose a gestural origin for language. Unfortunately, there is no evidence that sign languages are universal in human society, as speech grammars are. Moreover, there are no vestiges of a once-dominant universal sign language, as there should be if it was the original form of language. In my view, the fact that sign languages are possible does not suggest a gestural Ur-tongue for humans; rather it suggests that mimesis came first, as a whole-body skill, and that many linguistic constructions are mapped directly on to mimetic ones.

The universal form of language is undoubtedly speech. Language is thus normally layered on top of a mimetically-skilled phonological system during its development. Phonology has the special virtue of being able to generate a virtually infinite number of easily retrievable sound-patterns for symbolic use, to form the articulatory gestures that make up the phonological components of language. It provides the ideal vehicle for morphological experimentation. Human retrieval capacity for oral words is extraordinary—we carry around tens of thousands, and in the case of some multilinguals, hundreds of thousands, of words; in the wild, most other species, from bees to the great apes, seem to be limited to at most a few dozen expressions; and this limitation even applies to Cheney and Seyfarth's (1990) vervet monkeys. The only relatively close parallel to human phonological skill might be the vocal skill of certain birds, such as parrots and mynah birds. However, their adaptations were specialized at all levels—confined to sound, and to a very limited arena of communication.

In contrast, the human vocal adaptation came as part of a domain-general cognitive adaptation. This is evident in the fact that language is not confined to the vocal modality. Human mimesis is inherently supra-modal, cutting across all voluntary motor systems. Thus when phonology malfunctions in development, other mimetic subsystems may be harnessed by the language system, as they are in the deaf. This was shown most clearly in Petitto and Marentette's (1991) study of deaf infants' hand-babbling in sign-language environments, which occurs at exactly the same time as phonological babbling, and has the same properties. Deaf infants, growing up in deaf-signing households, are very good at miming the motor *principle* behind signing, if not the specific signs themselves. In this case, the motor principle is the same, whether vocal or manual; to produce actions repetitively at a certain rate, that most resemble those that they see others producing. Thus their manual 'babbling' reflects their expressive environments accurately, on a purely mimetic level. This is exactly

what babbling infants do in hearing households—they create models, in their actions, of one of the most salient dimensions of motor behaviour observed in their families—repetitive phonological acts.

Babbling, whether oral or manual, is reference-free in the linguistic sense—that is, it has no linguistic meaning—but it is truly representational, in that babbling patterns are (eventually) excellent motor-models of the expressive patterns infants observe around them. These models are *morphological*, but not yet linguistic, because they are reference-free and grammar-free. They reproduce not only the elementary morphemes of language, but also the larger mimetic envelope of expression—for instance prosody, and the habit of alternation, or ‘waiting one’s turn’ in expressive exchanges. Since babbling is free of a linguistic frame of reference, the brain mechanism that supports it does not have to be linked to language per se; rather, these eight- to ten-month old infants look like pretty good mime artists. The supra-modal nature of their babbling is very revealing—the fact that babbling is not confined to phonology is consistent with the idea that a domain-general mimetic adaptation evolved first, serving as the initial morphological foundation for language evolution. High-speed phonology seems to have developed later, as a specialized mimetic subsystem specifically driven by selection pressures that favoured the use of sound as the major modality for language. This is not to say that ontogeny recapitulates phylogeny; the language of infants would not necessarily resemble the forms that early hominid communication took. But ontogeny can reveal something about the functional hierarchies underlying language (see Nelson 1996, and my comment, Donald 1994, on Karmiloff-Smith’s 1992 book *Beyond Modularity*). Skills are nested in contingent hierarchies and, in language, a capacity for morphological modelling precedes the emergence of lexical skill.

### **The other side of the lexical entry—lemmas**

Language-first evolutionary theorists might try to hold the line at this point by conceding that, yes, basic motor skill advances might have been necessary for protolanguage, but were probably driven by the same selection pressures that produced protolanguage, and indeed, were an integral component of it. Again, I disagree.

In protolanguage, as in fully-developed language, the construction of a lexical entry requires more than morphophonology; it also requires the other side of the lexical entry, the ‘lemma’ (Levelt 1989). This necessitates a second level of mental modelling, one that harnesses the mind’s underlying morphological skills to construct linguistic statements and propositions. This second level—lexical semantics and grammar—is absent in mimesis. The integration of morphophonological addresses into a larger descriptive system is an inherently social activity, and one is tempted to predict that this process could not be confined within the isolated brain; that is, one should not expect to find the ‘language acquisition device’ that Chomsky (1965, 1988) predicted entirely inside the individual brain. Rather, the emergence of language depends on a community of brains in interaction. There is not yet any viable computational model of this process, and neural network models have not yet reached the point where anything so complex could be simulated. The process of mapping the ‘lemma’ or

meaning-based side of the lexicon on to the form of the symbol—whether it is phonological or manual—involves much more than the association of a discrete form with a discrete meaning. Phonology, like any mimetic system, works according to a metaphorical principle, but the lemma involves incredibly subtle forms of meaning that cannot simply be reduced to other symbols, if Wittgenstein (1922) or Johnson-Laird (1983) are to be believed. Word-forms and meanings both tend to be fuzzy, and neither side of the lexical entry is clearly defined or discrete. The tension between word-form and meaning is a creative one that greatly increases the range of things that can be represented.

The shift here is very fundamental—mimetic models generally work on an *analogue* principle, whereas linguistic models work on the principle of segmentation, labelling and thus the *consensual definition* of reality. Where mimetic models are driven by perceptual similarity and metaphor, linguistic ones are driven by definable, specifiable, and arbitrary relationships. Mimesis is truly representational on its own level; it underlies gesture, mime, body-language, dance, ritual, some music, and most non-verbal communication. But language captures reality in a different way; and linguistic knowledge emerges as an independent level of representation. The underlying envelope of speech-acts (see Fetzer 1993, and my reply at the end of the same article) is thus mimetic, not linguistic; and linguistic principles do not need to be introduced until some subset of mimetic inventions—call it the proto-lexicon—acquires a different cognitive role from mimesis.

The shift from mimesis towards protolanguage, which we can see in two-year-old children, is therefore a shift from motor homology and perceptual metaphor towards analysis and differentiation. Instead of being directly driven by perceptual metaphor, as mime is, words differentiate and specify the components of perceived relationships. Parsing the environment, that is, assigning labels to various 'views' of perceived events, is central to language. A linguistically-able mind must be able even to parse its own representations, and those of others.

The impact of this second level of skill on manifest group behaviour again changes the cultural picture of humanity. Able to perform such 'parsing,' the species produced a new, and very different, pattern of collective representations. It is still an open question whether grammar, and metalinguistic skills at the level of discourse, require a separate adaptation in addition to phonology and lexical invention. If one were to try to meet all of Fodor's (1983) requirements for a true linguistic 'module', a separate grammar module would surely fail, inasmuch as it seems to be completely interpenetrable with the rest of language, and closely tied to semantics. The neuropsychological case for a separate grammar module is further weakened by cross-linguistic studies of aphasia, which show that no specific brain lesion, nor specific pattern of grammatical deficit, is found in agrammatics of all languages (Bates *et al.* 1987). The whole perisylvian region of the left hemisphere seems to be dedicated to language, with function words and grammatical rules being stored in the same tissue as other aspects of the lexicon. Grammars are thus better regarded as a natural by-product of the process of parsing and definition; I see no compelling need to place grammatical invention in a separate category.

The invention of a symbol is a reciprocal process—form is mapped on to meaning, while meaning is *defined* by that same process, in a reciprocal tension that is revealed



by the fact that language is never static. This tension is still evident, after at least 45 000 years of lexical invention. Languages are constantly changing their particular mappings of form on to meaning; for instance the entire Indo-European group of languages, including languages as diverse as Sanskrit, Gaelic, Latin, and Greek, seems to have evolved from a common ancestor within the past 7000 years (Renfrew 1987). This pattern of incessant innovation suggests that the cognitive process driving lexical invention—the need to define and redefine our maps of meaning on to form—is more fundamental and considerably less rigid than the specific forms and rules of language at any given moment.

In conclusion, hominids could not have evolved a capacity for 'protolanguage' without meeting at least two preconditions. The first was the prior evolution of a more powerful central motor capacity that gave hominids a degree of voluntary movement control that does not exist in any other species. This new motor-modelling skill, or 'mimesis', allowed them to vary and elaborate upon the entire voluntary motor repertoire of the primate line. This provided a capacity for considerable morphological invention, and thus, in theory, some of the morphophonetic capacities underlying language must have been in place first. However, their communicative use at this preliminary stage was probably restricted to action-metaphor, mime, and rudimentary gesture. On the present evidence, general mimetic skill would seem to have sufficed for the achievements of archaic *Homo* and could have served as the basis for shared expressive custom, leaving the human line with a legacy of non-verbal culture both in infancy, where language is acquired in a mimetic context; and in adults, where forms like dance, pantomime, and a public 'theatre' of expression still play a very major role, and remain largely independent of language in both their social function and guiding principles. This non-verbal background of custom and expression eventually set the cultural stage for protolanguage, by establishing a public arena for non-verbal representation, in which a disambiguating device would have been found useful.

The second precondition was the evolution of a more abstract capacity for expressive modelling, one that did not rely entirely on perceptual metaphor as the carrier of conventional meaning. This capacity, apparently unique to sapient humans, 'defines' and parses reality by employing specific classes of morphological invention as explicit memory labels; note that the evolution of this second capacity therefore depends upon the pre-existence of the first. In evolution, this second step provided the cognitive mechanism needed to construct the other side of the lexical entry, or lemma, and elevated the morphological inventiveness of hominids to a truly linguistic level. This innovation, probably first found in nascent form in early sapients, would have triggered the structural basis for evolving a capacity for lexical invention and the first protolanguages.

A corollary of this hypothesis is that the evolutionary origins of the elaborate mimetically-driven systems of expression common to all human cultures are distinct from the roots of language per se. Protolanguage, even in modern human infants, floats on the surface of such a system (Nelson 1996), and the latter must be at least partly in place before the child can begin to invent linguistic descriptions. Similarly, in its first appearance in evolution, protolanguage must have emerged slowly from a cultural environment ruled by mimetic principles—that is, by action-metaphor.

## References

- Armstrong, D., Stokoe, W., and Wilcox, S. E. (1995). *Gesture and the nature of language*. Cambridge University Press, Cambridge, MA.
- Bates, E., Friederici, A., and Wulfeck, B. (1987). Grammatical morphology in aphasia: evidence from three languages. *Cortex*, **23**, 545–74.
- Bickerton, D. (1990). *Language and species*. University of Chicago Press.
- Bickerton, D. (1993). Putting cognitive carts before linguistic horses. *Behavioral and Brain Sciences*, **16**, 749–50.
- Boesch, C. and Boesch-Ackermann, H. (1991). Dim forest, bright chimps. *Natural History*, **9**, 50–7.
- Brownman, C. P. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, **6**, 201–51.
- Calvin, W. (1983). A stone's throw and its launch window: timing precision and its implications for language and the hominid brain. *Journal of Theoretical Biology*, **104**, 121–35.
- Cheney, D. and Seyfarth, R. (1990). *How monkeys see the world*. University of Chicago Press.
- Chomsky, N. (1965). *Aspects of a theory of syntax*. MIT Press, Cambridge, MA.
- Chomsky, N. (1988). *Problems of knowledge*. MIT Press, Cambridge, MA.
- Corballis, M. C. (1989). Laterality and human evolution. *Psychological Review*, **96**, 492–505.
- Corballis, M. C. (1991). *The lopsided ape: evolution of the generative mind*. Oxford University Press, New York.
- Darwin, C. (1871). *The descent of man*. John Murray, London.
- Dennett, D. (1992). The role of language in intelligence. *Darwin Lecture*, Cambridge University, March 6, 1992. Publication CCS-92-3, Center for Cognitive Studies, Tufts University, Boston, MA.
- Donald, M. (1991). *Origins of the modern mind: three stages in the evolution of culture and cognition*. Harvard University Press, Cambridge, MA.
- Donald, M. (1993a). Précis of 'Origins of the modern mind', with multiple reviews and author's response. *Behavioral and Brain Sciences*, **16**, 737–91.
- Donald, M. (1993b). Human cognitive evolution: what we were, what we are becoming. *Social Research*, **60**, 143–70.
- Donald, M. (1993c). Hominid enculturation and cognitive evolution. *Archaeological Review from Cambridge*, **12**, 5–24.
- Donald, M. (1994). Representation: ontogenesis and phylogenesis. *Behavioral and Brain Sciences*, **17**, 714–715.
- Donald, M. (1995). The neurobiology of human consciousness: an evolutionary approach. *Neuropsychologia*, **33**, 1087–102.
- Donald, M. (1996). The role of vocalization, memory retrieval, and external symbols in cognitive evolution. *Behavioral and Brain Sciences*, **19**, 159–64.
- Donald, M. (1997). The mind from a historical perspective: human cognitive phylogenesis and the possibility of continuing cognitive evolution. In *The future of the cognitive revolution* (ed. D. Johnson and C. Ermeling), pp. 478–92. Oxford University Press.
- Donald, M. (1998). Mimesis and the executive suite: missing links in language evolution. In *Approaches to the evolution of language: social and cognitive bases* (ed. J. Hurford, M. Studdert-Kennedy, and C. Knight), pp. 44–67. Cambridge University Press.
- Fetzer, J. (1993). Evolution needs a modern theory of the mind. *Behavioral and Brain Sciences*, **16**, 759–60.
- Fodor, J. A. (1983). *The modularity of mind*. MIT Press, Cambridge, MA.
- Gamble, C. (1993). *Timewalkers: the prehistory of global colonization*. Harvard University Press, Cambridge, MA.
- Greenfield, P. M. (1991). Language, tools and the brain: the ontogeny and phylogeny of hierarchically organized sequential behaviour. *Behavioral and Brain Sciences*, **14**, 531–95.

- Greenfield, P. M. and Savage-Rumbaugh, E. S. (1990). Grammatical combination in *Pan paniscus*: processes of learning and invention in the evolution and development of language. In *Language and intelligence in monkeys and apes* (ed. S. T. Parker and K. R. Gibson), pp. 540–78. Cambridge University Press.
- Johnson-Laird, P. N. (1983). *Mental models*. Harvard University Press, Cambridge, MA.
- Karmiloff-Smith, A. (1992). *Beyond modularity: a developmental perspective on cognitive science*. MIT Press, Cambridge, MA.
- Kosslyn, S. (1988). Aspects of a cognitive neuroscience of imagery. *Science*, **240**, 1621–6.
- Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. MIT Press, Cambridge, MA.
- Lieberman, P. (1984). *The biology and evolution of language*. Harvard University Press, Cambridge, MA.
- Mithen, S. (1996). *The prehistory of the mind*. Thames and Hudson, London.
- Nelson, K. (1996). *Language in cognitive development: emergence of the mediated mind*. Cambridge University Press.
- Paivio, A. (1991). *Images in mind: evolution of a theory*. Harvester Wheatsheaf, Toronto.
- Parker, S. T. and Gibson, K. R. (1979). A developmental model for the evolution of language and intelligence in early hominids. *Behavioral and Brain Sciences*, **2**, 367–408.
- Petitto, L. and Marentette, P. F. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science*, **125**, 1493–6.
- Pinker, S. (1994). *The language instinct: how the mind creates language*. Morrow, New York.
- Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, **13**, 707–84.
- Renfrew, C. (1987). *Archaeology and language: the puzzle of Indo-European origins*. Jonathan Cape, London.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Braake, K. E., Williams, S. L., and Rumbaugh, D. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*.
- Tooby, J. and Cosmides, L. (1989). Evolutionary psychology and the generation of culture. Part I. *Ethology and Sociobiology*, **10**, 29–49.
- Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., Sevcik, R. A., and Rumbaugh, D. A. (1993). Pan the tool-maker: investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological Science*, **20**, 81–91.
- Wittgenstein, L. (1922). *Tractatus logico-philosophicus* (trans. D. F. Pears and B. V. McGuiness, 1961). Routledge and Kegan Paul, London.