Edited by James R. Hurford, Michael Studdert-Kennedy and Chris Knight

Approaches to the Evolution of Language

Social and Cognitive Bases

4 Mimesis and the Executive Suite: missing links in language evolution
Merlin Donald 44

CAMBRIDGE UNIVERSITY PRESS 1998
List of contributors vii
Acknowledgments ix

Introduction: new approaches to language evolution
Michael Studdert-Kennedy, Chris Knight and James R. Hurford 1

Part I  Grounding language function in social cognition
1  Introduction: grounding language function in social cognition
   Chris Knight 9
2  On discontinuing the continuity–discontinuity debate
   Jean Aitchison 17
3  The origin of language and cognition
   Ib Ulbaek 30
4  Mimesis and the Executive Suite: missing links in language evolution
   Merlin Donald 44
5  Ritual/speech coevolution: a solution to the problem of deception
   Chris Knight 68
6  Theory of mind and the evolution of language
   Robin Dunbar 92
7  Old wives' tales: the gossip hypothesis and the reliability of cheap signals
   Camilla Power 111
8  Altruism, status and the origin of relevance
   Jean-Louis Dessalles 130
9  The evolution of language from social intelligence
   Robert Worden 148

Part II  The emergence of phonology
10 Introduction: the emergence of phonology
    Michael Studdert-Kennedy 169
11 Long-call structure in apes as a possible precursor for language
    Mária Ujhelyi 177
12 Social sound-making as a precursor to spoken language
    John L. Locke 190
13 The particulate origins of language generativity: from syllable to gesture
    Michael Studdert-Kennedy 202
14 Evolution of the mechanisms of language output:
    comparative neurobiology of vocal and manual communication
    Peter F. MacNeilage 222
15 Systemic constraints and adaptive change in the formation of sound structure
    Björn Lindblom 242
Mimesis and the Executive Suite: missing links in language evolution

MERLIN DONALD

There is considerable disagreement about the nature of the deep neurocognitive systems that generate human language, and how they might have evolved. The main bone of contention is whether the underlying generator of human language is continuous with the machinery of the primate brain, or something completely new. There are two diametrically opposed schools of thought on this issue. The Cartesian school insists on a special status for language, and adopts what might be called a traditional noncontinuity stance. To Cartesianists, language is a truly novel adaptation, irreducible and unique. Consequently the language generator must be a qualitatively new device, and the governing mechanisms of language must be fundamentally different from the neurocognitive machinery that preceded them in evolution. These a priori dicta are supported by the fact that our closest relatives, the apes, are quite helpless in the distinctively linguistic realms of story-telling and narrative description, and remain so even when they are given a lexicon of symbols by human trainers. Their linguistic helplessness contrasts dramatically with the extraordinary linguistic talent of human children. This suggests a deep cognitive divide between apes and humans.

The second school of thought, which might be called the neurobiological approach, places continuity first, and remains sceptical about the special status of language. It affirms that natural selection, like diplomacy, must practise the art of the possible, and can select only from raw material already in existence. Pre-existing primate capacities must have provided the evolutionary wedge or ‘foot in the door’ that culminated in human language. There might have been a dramatic discontinuity of function in the evolution of language, but there could not have been any discontinuities of mechanism. A parallel might be found in the emergence of flight in birds, which is also a major functional
discontinuity in evolution. The continuity principle is not threatened in this case, because the mechanisms of flight in birds are continuous with the mechanisms of terrestrial locomotion in their reptilian predecessors. There is no compelling reason to expect that other functional discontinuities, including language, are different in this regard.

Cartesians might reply that continuity theory cannot account for the specialness of human language. That criticism had some validity in the past. Behaviouristic theories of language were not successful. But, perhaps because of this highly visible failure, many modern cognitive scientists accept too easily the necessity of a discontinuous language acquisition device. This acceptance is unproductive, because objections levelled at behaviourism do not necessarily apply to a more inclusive neurocognitive approach to language. Unless we are willing to abandon the theory of evolution altogether (as some diehard Cartesians apparently are) we must maintain the continuity principle. This imposes a heavy obligation on theorists in this field to build credible bridges between primate cognition and human language, and to purge any implied saltation or evolutionary miracle from their proposals.

1 Modular versus domain-general models of language evolution

In addition to the difficult question of establishing continuity, language evolution involves an architectural issue: modularity. The architecture of the mind has sometimes been presented as a system of interconnected modules, each of which performs a specific cognitive operation. Modules are relatively autonomous brain systems that gather and analyse input, resolve patterns, store memories and generate specific classes of action. These systems are, by definition, specialists, dedicated to one cognitive function. In contrast, the supposedly nonmodular components of mind are generalists, open-ended, amodal and concerned primarily with co-ordinating the activity of the whole organism. Fodor (1983) and many others have called the nonmodular region of the mind the ‘central processor’. Baars (1988) has called it the ‘global workspace’. Others have labelled the same central processor the ‘supervisory system’ of the brain, and others have referred to ‘domain-general’ central processes. Whatever they may be called, these nonmodular components of mind draw on data generated by various modular specialists, creating an overall response, plan, or strategy. I prefer to call these processes, which constitute a cluster of functional subsystems, the executive processes of the mind, because they are in a position of control. In any architectural diagram of the mind, they are typically placed at the peak of the cognitive hierarchy.

It may seem obvious that language is too broad a process, involving too many subsystems, to be considered a module in Fodor’s sense. But this is not necessarily the case. Modules are not necessarily dedicated to narrow functions. There are gradations of modularity, and some larger modules of the brain are dedicated to quite complex cognitive functions that employ several lower-level modules to their ends. A good example of this is a specialized primate cortical subsystem that relays retinal disparity information (which gives good stereo vision within a three-foot radius) to the frontal premotor centres for fine hand control. This is evidently a complex module that supports the unique visual–manual skills of primates. It is a fairly open system, but still basically encapsulated, and dedicated to one special function. Thus, it meets the definition of a bona fide module. It is probably typical of the more complex cognitive modules in the nervous system in that it feeds into, and receives inputs from, a number of high-level modules related to vision and hand control.

Assuming that language could somehow be detached from other aspects of higher cognition, the possibility of a similar high-level modular subsystem for language cannot be ruled out in principle.

The Cartesian approach leads inevitably to a modular evolutionary scenario for language. Cartesians insist that scenarios of language evolution must, at the very least, account for the existence of a unique human grammar module that, by its very nature, would be placed quite high in the architecture of mind, and depend upon a number of lower-level modules for its function. Both Bickerton (1990, this volume) and Pinker & Bloom (1990) have argued that the basic elements of Universal Grammar cannot be accounted for by any known feature of human serial motor behaviour, and must therefore be produced by a preprogrammed language module built right into the human brain. They seek empirical proof of their theories in ‘language genes’ and dedicated brain structures, such as the so-called language regions of Broca and Wernicke.

Although language capacity must obviously reside in some feature of the human brain, the search for a fixed, localizable brain module for language has been a failure. After a century of searching, there is still no identifiable cortical subsystem, no special Brodmann area, no
commissure or fasciculus, no characteristic laminar organization, no distinctive cellular or columnar structure, no special linguistic neurotransmitter system, indeed no known feature of the cortex or any other part of the brain that can be identified universally and exclusively with language. Moreover, the patterning of language breakdown in neurological cases, especially in those with syntactic deficits, does not support a straightforward modular solution for grammar. The genetic evidence for a language or grammar module is even more problematic, given the enormous epigenetic plasticity of the human neocortex during language development.

The most common neurobiological alternative to a Cartesian module for language takes the form of a quasi-modular, rather than a nonmodular, solution. The mainstream form of such a proposal is that language originated in a mechanism that was not initially dedicated to language per se, and is thus somewhat ‘leaky’, and domain-general, but nevertheless specialized for communication. Many of these proposals are variants on the motor-origins theory of language evolution. The motor-origins theory proposes that the evolutionary adaptation for language is really the end-result of a series of evolutionary modifications to the serial motor control systems of the primate brain. The strongest evidence in support of this idea is the close association between human praxis and language, which is evident in the genetic linkage between handedness and language dominance, and in the clinical correlation between aphasia and apraxia. Lieberman (1973, 1975, 1984), Kimura (1976, 1979) and Calvin (1993) have suggested that human language might have originated in the serial-ordering capabilities of the primate motor system, coming under intentional control in hominids, and eventually generalizing this property to a more recently evolved hominid vocomotor system. There might also be a link between gesture and language origins; this position has been expanded and articulated in detail by Armstrong, Stokoe & Wilcox (1995), who argue that, at the governing level, language is a nonmodular adaptation that originated in the human capacity for gesture.

Corballis (1989, 1991) and Deacon (1988, 1995) are in basic agreement with the motor-origins approach, but they have added a computational twist to the theory: namely, that a common property, ‘generativity’, underlies both praxis and speech. Generativity is the ability to build an infinity of recombinatorial sequences out of a finite set of discrete components. Greenfield (1991), Calvin (1983), MacNeillage (1992) and

Studdert-Kennedy (this volume) have all suggested closely related variations on that central idea, trying to bridge the gap between primate and human expressive skills in terms of changes to the computational principles governing the movement system. Although these proposals do not necessarily insist on the emergence of a discrete language module, they do propose a left-hemisphere motor-sequencing adaptation for the special kind of recombinatorial control that supports both praxis and language. All these theorists adhere to the continuity principle.

A second class of quasi-modular proposal has ensued from the literature on social cognition. Tomasello, Kruger & Ratner (1993) have proposed that language could only have emerged after hominids had evolved the social skills that are evident very early in human development, and that are normally the precursors of language. These social skills include a capacity for mindreading – that is, for understanding the intentionality of other minds – which forms the basis for the later development of both imitation and intentional expression. Thus the origins of language are tied to expanded conceptual capacity, rather than motor skill per se. Tomasello et al. concede something to the motor-origins theories inasmuch as they believe that the human capacity for imitation is important in evolving language; but even in the case of imitation, they see the primary cognitive adaptation as conceptual, rather than motoric, in nature. In the view of these authors, people can only imitate the actions of others if they can grasp their intentions; this allows them to unify conceptually an otherwise perceptually complex and confusing episode. Dunbar (1993) and Aiello (1995) have also supported the importance of social cognition in language evolution. They have suggested that language emerged as part of a more inclusive hominid adaptation for sustaining larger social groups. Their theory, vague about cognitive details, seems to place language squarely in the realm of a domain-general conceptual skill, or a communicative skill with strong ties to social cognition.

Most authors in this field, including myself, will concede that language has some degree of modularity in its evolution. The human vocal apparatus, and the extraordinary capacity for high-speed phonology associated with it, are obviously based on a modular adaptation. Human vocal skill represents a significant break with the primate pattern, and in itself constitutes a major functional discontinuity in primate evolution, as well as an empirically verifiable new module. Humans simultaneously evolved a corresponding phonetic memory capacity for sound-based
lexical addresses. But, granting that vocalization and its spinoffs are regarded as an essential part of the evolutionary emergence of language, they must be seen as a side-story to the main evolutionary question, because vocal phonology is neither necessary nor sufficient for language. The primary evolutionary story centres on representation itself, and it is not at all clear that representational skill is modular in the sense of being restricted to its use in language.

I have proposed a synthetic approach (Donald, 1991, 1993a, 1993b, 1993c, 1995, 1997) in which both elements, social cognition and the need for improved praxis, play a role in evolving a cognitive capacity that I have called mimesis. Mimesis is a nonverbal representational skill rooted in kinematic imagination — that is, in an ability to model the whole body, including all its voluntary action-systems, in threedimensional space. This ability underlies a variety of distinctively human capabilities, including imitation, pantomime, iconic gesture, imaginative play and the rehearsal of skills. My hypothesis is that mimesis led to the first fully intentional representations early in hominid evolution, and set the stage for the later evolution of language. The emergence of mimetic capacity involved changes that were primarily in the brain's domain-general or executive capacities. Mimesis, and language itself, are therefore traceable to a modification of the hominid executive brain, or Central Processor. This chapter will expand on that idea.

2 An additional consideration: language does not self-install

One additional point, and one that has generally been neglected, is relevant here. The language brain does not act like a preprogrammed capacity in one very important way. Even when it is given rich material on which to exercise its powers, it does not automatically self-install or self-trigger. This can be seen in people who are linguistically isolated from birth, such as the nonsigning deaf. They do not invent languages, even when they are exposed to an environment rich with describable experiences. One might expect that, given an event-structure to describe, a preprogrammed, fully equipped language acquisition device would dive into the task of labelling and describing the world. This would at least afford the individual brain some internal means of classifying the environment, and perhaps even of thinking symbolically. Such a capacity would undoubtedly have great adaptive advantage.

But the brain does not do this, ever. Without extensive cultural programming, the mind does not suspect the possibility of language, not even of inner language. The supposedly automatic process of language assembly is therefore not at all automatic; it depends heavily on extensive social interaction. It is hard to imagine how a fully preprogrammed language module could exist in such a form that it would not spontaneously self-trigger simply by being exposed to the episodic event-structure of the world, just as the visual system triggers vision, simply by being exposed to normal levels of patterned light (curiously, mindreading skill self-triggers in the nonsigning deaf, and seems more easily defensible as a module than language itself).

Perhaps we have not faced the neurocognitive consequences of Saussure’s famous assertion that languages are inherently collective and consensual communication systems. Languages apparently have no independent possibility of existence in individual brains. Single brains simply do not invent languages or symbols, and in isolation do not appear to have the capacity to do so.

What, then, is language capacity? Where might it reside? If it resists localization, and fails to self-trigger during development, then it must be assembled in the brain very differently from the way sensory-motor modules are assembled. It depends on complex epigenetic programming that does not take place without intensive cultural interaction. This implies that much of the replicative information needed to perpetuate language is stored in culture, not in the genes. While the human brain obviously has a capacity for language in a general sense, it might take the form of a broad potential for communication and representation, rather than a preset language system with precise specifications, such as Universal Grammar, built right in. The question then becomes an epigenetic, rather than a genetic one. The central issue is to determine the nature of the brain’s interaction with the communicative environment. How does the individual brain ‘couple’ with culture to create language, and how could such a process have begun?

3 Cognitive fundamentals and the role of the executive brain

The idea that the evolution of language was heavily dependent on the expansion of a domain-general executive brain may appear odd at first,
especially to linguists. But the executive brain makes possible many of the complex cognitive functions that exist not only in humans, but also in other mammalian species. It provides the cognitive fundamentals of most higher mental activity. These fundamentals allow the individual brain to ‘couple’ with the labyrinthine representational mazes that constitute the building-blocks of culture. Cognitive engagement with culture stands at the heart of our uniquely human forms of representation, and the executive brain directs this process.

The human executive brain thus provides the basis for cultural engagement, and, indeed, of cultural invention in the first place. It is fundamental to all our unique communication skills, including language. It is not simple in structure. In fact, the executive brain is an extensive complex of brain subsystems that collectively supports a class of domain-general capacities, including various kinds of working memory, at least two or three different kinds of attentional selection, the basis for autoecuing or self-triggered recall, abstract self-representation, and much of the basis for social cognition. These skills were the evolutionary wedge from which language was fashioned by selection pressure, because of their close relationship to primate executive skills.

Before developing this notion any further, some common misconceptions must be corrected. One objection made by Tooby & Cosmides (1989) to the possibility of general-purpose adaptations in cognitive evolution is that it is difficult to imagine what kind of selection pressures could cause such (presumably) diffuse adaptations in any given species, given that the gene-pool is always responding to very specific environments. But this objection is based on a complete misunderstanding of what is involved in a domain-general cognitive adaptation. Domain-general aspects of the mind and nervous system are not unstructured or unspecialized in their evolutionary origins and biological function. They are not general-purpose, or diffuse, in that sense. On the contrary, domain-general executive skills are a specialization in themselves, like any other. There are many potential forms that a domain-general evolutionary change might have taken. The particular solution in the case of humans was successful because it responded directly to the specific selection pressures that were at work in human evolution. Even though, in architectural terms, this solution was domain-general, it was not ‘general-purpose’ in terms of adaptation.

Even from the more restricted perspective of cognitive architecture, the executive systems of the brain are not necessarily as amorphous as Fodor (1983) suggested in his important treatise on modularity. Executive systems perform very specialized functions, such as retrieving and holding relevant knowledge in working memory, switching between ongoing tasks, temporarily interconnecting various cognitive subsystems (as in many multimodal tasks), building and updating an ongoing model of the current environment, initiating searches of long-term memory, maintaining a hierarchy of priorities and so on. The elaborate component structure of the executive brain reflects that complexity.

The mammalian executive brain has a traceable evolutionary history. The gradual elaboration of executive brain functions is one of the signal features of mammalian evolution. This elaboration was related to the emergence of a unifying cortical mantle and the subsequent need to manage the traffic of the emerging mammalian central nervous system. The executive brain in humans is simply a modification, a particular variation, of the ancient mammalian core brain system that includes (in humans) much of prefrontal cortex, the tertiary areas of the parietal-temporal cortex and most of the insula, cingulate gyrus and hippocampus, in addition to many subcortical structures, including parts of the midline thalamus and basal ganglia. All these structures are known to play an important role in working memory, attention, voluntary movement, explicit recall, social cognition, planning and conscious function in general. They work together in the regulation of what is usually called controlled cognitive function. Relative to the rest of the brain, the relevant brain areas are larger in humans than in apes. Thus, it must be concluded that, during human evolution, these specific brain regions expanded much more than other neural structures, such as the brainstem or the primary sensory areas.

The executive brain is necessary, and complementary, to the specialized modules it helps to regulate. It is consistent with a general pattern of neural adaptation to the escalating unpredictability of the mammalian world, as mammals evolved a greater capacity for learning, thus leading to greater ontogenetic variability. As selection pressures moved hominids toward social complexity, their societies created even higher degrees of unpredictability that had to be dealt with in terms of memory demands, divided attention and so on. By the very nature of its functions, the human executive system had to become increasingly domain-general in its architecture. As an evolutionary adaptation, however, it was highly specialized.
4 Where humans excel: the Executive Suite

The expansion of the hominin executive brain corresponded to a change in our core capacities. This change is evident in a variety of behavioural and cognitive domains for which humans are specialized. In this section I will describe a series of task-paradigms on which, relative to apes, humans excel. All these paradigms tap our executive skills directly. A mastery of language is not necessary for the performance of any of them. Apes, and sometimes monkeys and other mammals, have the ability to perform most of these paradigms either poorly, or to a limited degree. Differences between apes and humans on these tasks are usually quantitative, not qualitative. These paradigms were selected by researchers because they were on the very edge of what primates can achieve, and situated, as Vygotsky (1978) might have put it, in their ‘zone of proximal development’. It follows that these functions could have been subjected to natural selection early in hominid evolution.

Hominid executive skills, complex as they are, must have evolved almost simultaneously, forming what is sometimes called a suite of interrelated adaptations. These might therefore be called the Executive Suite.

4.1 Metacognition, or self-monitoring

Metacognition is the ability to monitor one’s own performance. It is essential in the rehearsal of skill, inasmuch as a degree of self-evaluation is demanded if performance is to be improved by repetition. It is also important in maintaining a communication linkage over time – that is, monitoring the success of one’s own performance as a communicator. Humans have a highly developed, and primates a limited, metacognitive capacity. Smith, Schull & Washburn (in press) have shown this with a behavioural test consisting of what they call the ‘bailout’ option. Animals are trained to respond to a stimulus with a ‘yes’, ‘no’, or ‘bailout’ option. The bailout option indicates that the animal wants to skip the present trial, and advance to the next one, thus avoiding punishment for a wrong answer. Monkeys can do this task. Thus they can, to some extent, monitor their own likelihood of success or failure on a given trial. Rats cannot perform this task,1 but there may be other mammals that can.

1 However, rats are apparently able to achieve a limited kind of self-report; see Beninger, Kendall & Vanderwolf (1974).

4.2 Imitation

Apes and monkeys are capable of some imitation, but Tomasello et al. (1993) have argued that the imitative skills of apes fall far short of those of humans. In particular, they are not as good as humans at precise, detailed reproduction of sequences. It is as if they were lacking in the essential attentional and/or mnemonic power to focus on, and remember, the specifics of the imitated sequence.

4.3 Multitasking (consciously managing two or more operations)

Savage-Rumbaugh and her colleagues (1991, 1993) have suggested that chimpanzees and bonobos have most of the cognitive components for language, with one major exception: multitasking, or the ability to consciously manage more than one thing at a time. Multitasking is common among humans, and essential to intentional communication, since speakers must keep track of both their own behaviour and that of their audience, splitting their attention between two or more channels.

4.4 Autocuing, and explicit memory

There is a great methodological divide between the literatures on human and animal memory (see Sherry & Schacter 1987). Animals, including apes, have a very limited capacity for recall. Recall that of most other mammals, the ape’s memory is best addressed passively, and this limits the kinds of symbols that apes can be taught to use (they are much better at using prepackaged symbol-boards than they are at signing, for example). Humans, in contrast, can easily self-trigger, or autocue, their memories. Explicit recall in humans depends heavily on what Karmiloff-Smith (1992) calls the ‘representational re-description’ or ‘explicitation’ of reality.2 This skill is fundamental; hominids could not have invented lexicons unless they already had some capacity for creating retrievable morphology, a precondition for the emergence of protolanguage.3 Once language had evolved, it greatly enriched our capacity for explicit recall at higher levels of representation.4

2 This traces back to Bartlett’s (1932) concept of ‘reconstructive’ memory and is also related to Edelman’s (1992) notion of ‘recategorial’ memory. In all three cases, memory is part, or a reflection, of the modelling intellect.

3 For a more detailed discussion of this topic, see Donald (in press).

4 See also Nelson (1990, 1996).
4.5 Self-reminding (mnemonic uses of action)
The acquisition and maintenance of a complex skill require an ability to repeat an action in order to self-remind, so that long sequences can be acquired and remembered. Human children often use their own actions as mnemonic devices, but apes do not, at least not systematically. Self-reminding is related to autocuing skill, as well as to metacognition. It is an intermediate step to long-range intentional guidance behaviour.

4.6 Self-recognition
Apes can recognize themselves in a mirror, whereas monkeys and other mammals cannot (Gallup 1970). Thus apes possess the capacity not only to construct an objectivized image of themselves in the environment, but also to consciously examine and recognize rotations and transformations of that image. Humans are generally much faster at self-recognition than apes, and this capacity is central to praxis. Actors and athletes use this kind of kinematic imagination, or self-visualization, in training.

4.7 Purposive rehearsal
One of the most essential components of advanced skill is the capacity to rehearse action purposively. This complex self-supervisory capacity is unique to humans. Human children endlessly rehearse and refine their own actions, sometimes without any obvious reinforcement; for example, they might practise throwing stones at a tree, or balancing on one foot. Although apes play repetitive games, they do not appear to spontaneously rehearse and refine their skills. This involves many of the skills listed above – self-monitoring, autocuing, visual and kinetic imagination, advanced self-representation, and so on – which exist in limited form in primates, and could have evolved as a result of the convergent evolutionary expansion of these other elements. Purposive rehearsal requires the individual brain to evaluate its own outputs and conceive of idealized actions as a template for judging outcomes. Rehearsal involves a form of intentional representation; the act is used to represent itself (a kind of self-demonstration). This ability seems to be based on implementable imagery.

4.8 Action metaphor (iconicity in expression)
The repetition of an action to communicate what the act naturally conveys is the most elementary kind of intentional iconicity. This might be called action metaphor, and constitutes a rudimentary form of gesture. There is no evidence for the existence of this kind of intentional

behaviour in apes, but there are some reports of pseudo-gestures, which means that they may be quite close to having gestural skill. Thus, iconic skill in primates, or in early predecessors of humans, such as the australopithecines, may have been subjected to selection pressure.

4.9 Mindreading (imagining and understanding intentionality in others)
This capacity is domain-general in the sense that it is abstract and amodal. Inasmuch as they are capable of deception, mindreading is already present in monkeys and apes. However, their mindreading skills are evidently not sufficient to support the spontaneous appearance of a conventional system of gestures (cf. Tomasello, in press). This is mainly due to the absence of the other cognitive fundamentals listed here.

4.10 Reciprocal intentionality (proto-pedagogy)
This is two-way mindreading, or what might better be called mind-sharing. It is fundamental to pedagogy, in which one person consciously regulates the learning process in another, while the learner tracks the teacher’s intent. The teacher–apprentice relationship is central to human culture. It involves not only understanding the intentionality of others, but also engaging in intentional exchanges, and understanding how the mind of the other has been affected by the communicative act. This capacity exists in people who lack language (see Schaller 1991) but seems to be absent in apes, although some signs of it are present in their subtle social event-perceptions. Their perceptions, however, are much better than their productions, and even extensive enculturation has failed to ignite much by way of spontaneous reciprocal intentionality.

4.11 Spontaneous symbolic intention
This has never been convincingly demonstrated in apes, even enculturated ones, although there have been a few reports of symbolic invention in the form of small variants on existing expressive custom (cf. Wrangham, McGrew, De Waal & Helme 1994). Human children, in contrast, are very expressive with gesture as early as one year of age. Gestural skill is domain-general inasmuch as it can harness any input or output modality.

5 Boesch & Boesch (1984) claim that a quasi-pedagogical relationship exists between mother and infant chimpanzees, with regard to learning how to use certain tools. This could be due to imitation or emulation by the infant, rather than systematic instruction by the mother. Nevertheless, it is close enough to basic pedagogy to provide something for natural selection to act on.
This brief summary makes no claim to be exhaustive. But it does show that there is a fairly broad class of executive skills that distinguish humans from apes. These functions are heavily dependent upon the executive brain; extensive experimental evidence on the properties of the executive system is summarized in Baars’ book (1988) on this subject.

The whole executive complex – both brain anatomy and the executive capacities themselves – evolved during hominid evolution, and this raises the question of their origin.

Could all this not have co-evolved with language? I regard this as highly unlikely, for the following simple reason: language is the most complex of human skills; the simpler ones probably came first. Even protolanguage could not have evolved before the central executive apparatus had reached a certain point. Intentional communication demands a great deal of executive management. The moment-to-moment difficulty of acquiring and managing language in the real world is very high, because the cultural environment that encompasses language is constantly changing, and in need of very rapid updating. Moreover, human speakers often carry out several complex operations at once, in several modalities, simultaneously maintaining parity with the often several recipients of their communications. All these achievements demand central attentional control to a degree that is unheard-of in other species.

Human language is characterized by improved access to memory. Speakers gain voluntary access to their own memory systems, a skill that is singularly powerful. But this is not unique to language; it applies to other special human abilities. For instance, an actor can carry a repertoire of literally hundreds of conventional gestures, postures, sounds, facial expressions and attitudes. This skill pales in comparison to language, however, where the explicit retrieval of words and idioms must be achieved much faster and more accurately; no small matter when a lexical search must separate out a single item from a repertoire of tens of thousands.

Thus the distinguishing cognitive features of language, when compared with other highly skilled tasks, are its complexity and speed, and its demands on attention and working memory. Language is really a gigantic meta-task, requiring the co-ordination of an entire hierarchy of subtasks and sub-subtasks, regulated from working memory. Speakers hold in their working memories a model of their own past communications, and a model of what other participants have already said, tagged as to source; otherwise a two- or three-person conversation would cease to make sense very soon after having begun. This multichannel updating skill cannot be modular in any meaningful sense of the term; it must remain open to all sorts of inputs, including nonlanguage ones. Moreover, the working model must be immediately available for constant updating; one might say that the model is the speaker’s effective working memory. It guides the overall contour of communication for considerable periods of time. This kind of model-building, including its role in working memory, is a major component of human executive capacity.

The inevitable other face of controlled performance is a capacity for automatization on a massive scale. The acquisition of any special human skill, including language, depends on executive control. But the efficient routine use of those same skills depends more on an ability to automatize them, leaving conscious capacity free to deal with more urgent matters, such as the novel, the complex and the unexpected. Thus the basic cognitive operations underlying language, such as lexical search, sentence assembly and the generation of linguistic morphology, must become automatized so that the speaker is free to take a complex communicative context into account. By increasing their executive capacity – including an expanded capacity for automatization – humans have effectively become self-programmable organisms.

5 Episodic competence in primates and humans

There were other demands placed on the executive brain in primate evolution, especially in the social domain. Acquired social skills are among the most esoteric achievements of primates. Remembering and responding to social situations is a complex multi-channel task that demands the integrated use of several basic brain functions. One of

6 The best evidence of automatization comes from the study of trade-off effects while doing several things at once, in which one task, say, driving a car, interferes with performance in another, say, carrying on a conversation in Italian. Inexperienced performers, for instance someone who is simultaneously learning to drive and to speak Italian, find it impossible to perform two non-automatized tasks at once. Once skills are sufficiently automatized, multitasking becomes possible without significant tradeoffs. Automatized skills can be nested in quite complex hierarchies, as in piano-playing, for instance.

7 The best summary of the evidence for these operations in language is found in Levelt (1989).
these functions is social event-perception, and its corollary, event-parsing, which includes an implicit understanding of the agents, their interactions, their effects on the contingencies of action, and the consequences of perceived episodes. These are the product of highly abstract amodal capacities.

I have called this sophisticated style of event-representation episodic cognition because it is concrete or episode-bound. Apes are among its most skilled practitioners, and are very clever problem-solvers within any brief experiential episode. Episodic skills sit at the top of the cognitive hierarchy. Episodic competence demands considerable executive skill, including an ability to attend selectively (the parsing of episodes is, after all, a matter of registering the event selectively and interpretatively), the capacity to sustain vigilance over fairly long periods, and a prioritizing capacity that allows strategic decisions as to the best course of action, when several courses are possible.

Above all, episodic cognition makes great demands on working memory. Various facts and implications that are inherent in social episodes must be kept in mind when negotiating, for instance, the subtleties of chimpanzee social behaviour. These might include one’s place in the dominance hierarchy, territorial interests, past interactions with other players and past interactions between the others involved in the episode, as well as one’s immediate objective (for instance to be groomed by a particular individual). These facts need constant updating, since social relationships can change suddenly.

Humans also apply their executive capacity in the regulation of social behaviour. In many ways, social behaviour is the most conscious of our behaviours, because it is unpredictable and complex, and thus requires constant updating and vigilance, careful attention to multiple simultaneous channels of input and a good memory for past social interactions. Above all, social behaviour involves the constant construction and updating of mental models of events and actions, a process that is highly conscious and controlled, precisely because it must remain well matched to a constantly changing social and cultural environment. In this, social behaviour is like many other human skills, including language.

Language is undoubtedly much more abstract in its structure than episodic cognition, and in human interactions it can dominate our thinking about social relationships to the point of obscuring the more basic episodic capacity that drives it. But it is important not to underestimate the complexity of episodic processes. Most importantly for evolutionary considerations, language – and intentional communication in general – seems to rely on the same kinds of executive skill as episodic cognition.

6 Mimesis: the linkage between primate event-cognition and intentional cultures

The primary reason for proposing an expansion of executive skills early in hominid evolution is that such a change would have maintained evolutionary continuity, while forming the basis of a cognitive adaptation that can account for the achievements of early Homo (this is where a close consideration of evidence from physical anthropology and archaeology is essential). The most credible form of early hominid culture, appearing with species Homo about 2 million years ago, is a mimetic style of thought and communication. Mimetic skill is the foundation for a variety of hominid achievements, including advances in toolmaking; co-ordinated collective endeavours such as fire use; sophisticated hunting and gathering; and the division of labour. This was a significant step away from the primate mainstream, without invoking anything so refined or powerful as language.

My hypothesis of the early evolution of mimesis achieves four important theoretical objectives. First, it accounts, as well as can be expected at this time-distance, for the archaeological evidence on the achievements of early hominids. Second, it explains the presence of a very salient (especially in the neuropsychological literature) analog mode of representation in modern humans, that represents our major alternative to language. Third, it accounts for a self-contained level of human culture that still operates on mimetic principles, and forms the basis for many expressive arts, such as dance and athletic display, pantomime, theatre, painting and many basic crafts.

Finally, it establishes the fundamentals of intentional expression in hominids, without which language would not have had an opportunity to evolve, since there would have been no existing communicative environment on which natural selection could act. There could not have been selection pressure to evolve such a sophisticated, high-speed communication system as modern language unless there was already a simpler, slower one in place. Mimesis fulfills that condition: it provides the preconditions out of which language, improbable an adaptation
though it may seem, might have evolved, simply as the need for disambiguation grew. Mimetic culture is considerably closer to language than primate expressive cultures.

The evolutionary process is not teleological; there could have been no foresight, no anticipation of language. This is also consistent with the early evolution and stability of mimetic culture. The archaic social adaptation of hominids was self-sufficient enough to last for a very long time. This is evident from the archaeological record. In both brain anatomy and associated cultural artifacts, *Homo erectus* was a rather stable species for well over a million years. There was clearly no need for it to evolve any further toward language to ensure its survival. To reiterate my rejection of teleology, there is no anticipatory element, implicit or explicit, in proposing the early evolution of mimesis. Mimesis would have been a successful adaptation in its own right, forming a satisfactory basis for the refinement and diffusion of praxic skills, for the development of much more intricate expressive scenarios and for the maintenance of some degree of social complexity.

Its early appearance as an archaic adaptation would not have ruled out further mimetic evolution at the time language emerged. Mimesis was, and remains, the basis for an analog style of communication and thought that is common in humans, but absent in apes. It is a highly sophisticated capacity, easily underestimated as to its importance in modern human behaviour. It would have made hominids formidable foes in any area in which competition was an issue, because the entire range of voluntary behaviours was brought under voluntary control.

Table 4.1 lists some common mimetic behaviours. They make up a loose representational hierarchy, including both the simpler forms of mimetic expression, such as pointing and intentional gaze, and the more complex forms, such as re-enactment and pantomime. The representational style of mimesis is variously describable as fuzzy, iconic and/or analog. It is consistent with what Burling (1993) has called an 'iconic' style of communication. The principle of similarity that links mimetic actions and their referents is perceptual, and the basic communicative device is best described as *implementable action metaphor*. Mimetic skill is a powerful device for communication: it can convey requests and commands, capture and hold the attention of others, show or declare, establish and maintain contact, refer explicitly to actions or events, demonstrate, oversee the actions of others and convey emotion. It also allows basic instinctual behaviours, for instance, pair-bonding and mating, to be made more elaborate and culturally specific in their forms.

In fact, the role of primate expressive instincts should not be overlooked in discussing human mimesis. It is highly likely that mimetic expression was initially built from preexisting primate emotional and social expressive algorithms. Humans have an unequalled expressive repertoire of limbic or involuntary gestures, body attitudes, facial expressions and sounds, including such uniquely human behaviours as laughter (itself a very powerful device for both bonding and social ostracism), extended and very intense grief reactions and many other universal behaviour patterns, which have been documented by Eibl-Eibenschütz (1989) and colleagues in cross-cultural studies. Such stereotyped emotional reactions and interactions project deep into the hominid past, and do not in themselves constitute mimesis. It is only when such reactions are used or elaborated intentionally that we have evidence of mimetic skill.

In addition to its role as a form of thought and representation, mimetic capacity serves a crucial subsidiary role in modern language, creating the communicative envelope of language, including such things as eye contact, facial expression, prosody and gesture, all of which
help in establishing and maintaining parity between various agents in a communicative setting. The close relationship of mimetic expression and culture to language development in modern human infants has been documented by Katherine Nelson in a recent book (Nelson 1996). Mimetic is the key to the earliest stages of human cultural engagement, and to the most basic repository of custom and many other forms of cultural memory. The common knowledge mediated by mimetic cognition eliminates the need for various kinds of explicit markers in oral communication. Additionally, many uses of language are basically mimetic, and in young children the use of spoken language is essentially secondary to their mimetic expressive agenda. Oral language typically emerges as an independent level of representation after four years of age (Nelson 1996).

Table 4.2 illustrates a hypothetical sequence in which mimetic communicative behaviours might have emerged. The most immediate fitness advantage for the evolution of hominid executive skills was probably a general improvement across the whole range of their voluntary repertoire, driven by intraspecific competition. A group of hominids with better communication and memory skills would have had a great advantage in most survival-related tasks, including toolmaking, hunting, scavenging and gathering.

The preexisting emotional expressions of hominids, including laughter, extended grief and other special facial, vocal and manual expressions, formed an important foundation for mimetic expression. The ease with which humans can employ these various response modalities, alone and in combination, testifies to the amodal position of mimesis in the communicative hierarchy. A domain-general capacity for modelling one’s own action is integral to mimetic and, by definition, whatever the initial reasons for its emergence might have been – whether better food-gathering, hunting, or toolmaking, or social innovations such as the division of labor – its emergence would automatically have generalized to better voluntary control over intentional expression in general. This has a variety of immediate social consequences: reciprocal gesturing, shared action-models, more elaborate games and signature behaviours that characterized specific groups. Even a small amount of intentional experimentation with these behaviours would have greatly complicated hominid social life. A small change in expressive ability would also have greatly increased opportunities for the use of deception, Machiavellianism, conspiracies and alliances, as well as inter- and intra-group competition. Mimesis would make social hierarchies more complex, and lead to greater differentiation of custom. With the arrival of mimesis, all the intricate social arrangements of primates were suddenly vulnerable to elaboration and differentiation. Natural selection had the opportunity to act on these, shaping a new form of hominid culture.

Mimetic vestiges persist in human affairs. Surely the most convincing evidence for the time-depth of the mimetic adaptation in hominids is the persistence of mimesis into the era of modern Homo sapiens. The vestiges of mimesis are especially evident in our irrepressible tendency toward conformity. A unique feature of human social life is an incessant pressure to conform to group norms, and to be accepted by one’s peer group. Another is the tremendously subtle cultural variations that occur outside the realm of language. The pressure to conform to those unwritten, unspoken cultural norms is virtually irresistible, whether it takes the form of imposed totalitarianism or the superficially voluntary conformity manifest in great public spectacles and demonstrations. Moreover, the threat of social ostracism is so profound that it can drive individuals to suicide. Mimesis is the driving force here; the pressure to conform reflects a profound instinct in humans to imitate the group, to absorb and internalize its cognitive models of reality.

Mimesis is thus as deeply instinctual as language itself. It is an adaptation in its own right, and the continuing importance of mimetic communication should not be underestimated. People can still be moved to act in a co-ordinated manner by a few gestures and utterances, and they cannot resist miming one another’s behaviour in large-group demonstrations. Language has only a minor role here. Mimetic culture has provided the cognitive foundation of human conventionality. It has

---

9 This type of theory is not necessarily recapitulatory in nature (cf. Nelson 1996; Donald 1995).
also provided humans with a public communicative space that is a necessary, if not sufficient, precondition for normal language acquisition. Most important, the fundamentals of articularatory gesture, from which all languages are built, were put in place when mimetic capacity emerged.

References


(1993b). Human cognitive evolution: what we were, what we are becoming. Social Research, 60, 143–170.


