



## Mirror neurons and the evolution of language

Michael C. Corballis

Department of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

### ARTICLE INFO

#### Article history:

Accepted 25 February 2009

Available online 1 April 2009

#### Keywords:

Mirror neurons

Language

Speech

Evolution

Gesture

### ABSTRACT

The mirror system provided a natural platform for the subsequent evolution of language. In nonhuman primates, the system provides for the understanding of biological action, and possibly for imitation, both prerequisites for language. I argue that language evolved from manual gestures, initially as a system of pantomime, but with gestures gradually “conventionalizing” to assume more symbolic form. The evolution of episodic memory and mental time travel, probably beginning with the genus *Homo* during the Pleistocene, created pressure for the system to “grammaticalize,” involving the increased vocabulary necessary to refer to episodes separated in time and place from the present, constructions such as tense to refer to time itself, and the generativity to construct future (and fictional) episodes. In parallel with grammaticalization, the language medium gradually incorporated facial and then vocal elements, culminating in autonomous speech (albeit accompanied still by manual gesture) in our own species, *Homo sapiens*.

© 2009 Elsevier Inc. All rights reserved.

### 1. Introduction

Ramachandran (2000) famously remarked that mirror neurons would do for psychology what DNA has done for biology—a remark that is in danger of being quoted almost as often as mirror neurons themselves are invoked.<sup>1</sup> Thus mirror neurons are said to provide an explanation for phenomena as diverse as imitation, action understanding, learnability, theory of mind, metaphor, and language. Failure of the mirror neuron system is also now widely accepted as an explanation for congenital neuropsychological deficits, such as autism (e.g., Baron-Cohen, 1995; Iacoboni & Mazziotta, 2007; Oberman et al., 2005). Ironically, though, mirror neurons were first discovered in the monkey brain, and monkeys are generally not credited with theory of mind, metaphor, or language—or autism.

The role of mirror neurons in imitation is more contentious. Although mirror neurons have not been recorded directly in humans, brain-imaging studies point to an equivalent system in the human brain, and this system is activated when people imitate action (Nishitani & Hari, 2000, 2002; Rizzolatti & Craighero, 2004). Yet monkeys appear to be incapable of imitation (Visalberghi & Fragaszy, 1990, 2002), suggesting that the mirror system did not evolve to mediate imitation. Rizzolatti and colleagues have suggested instead that the primary role of mirror neurons is in action understanding (Rizzolatti, Fogassi, & Gallese, 2001; Rizzolatti & Sinigaglia, 2008); that is, mirror neurons allow the monkey—or human—to understand actions performed by others by mapping those actions onto actions that it can itself perform, but they do

not mediate the actual imitation of those actions. As Hurford (2004) has pointed out, though, understanding in this sense need not imply the extraction of meaning, in the linguistic sense. To know what a given action means linguistically requires an extra step—or steps. For example, a person might easily copy the action of a sign in American Sign Language, without having any idea what it actually means in that language.

Nonhuman primates may not be totally incapable of imitation. For example, neonatal rhesus monkeys do imitate lip smacking and tongue protrusion performed by a human, but do not imitate mouth opening or hand opening (Ferrari et al., 2006). Mirror neurons may therefore elicit imitation of at least some simple actions already in the monkey's repertoire, although it may perhaps be questioned whether this is true imitation or effector priming. A further problem may be that monkeys do not naturally follow attentional cues provided by humans. Kumashiro et al. (2003) have shown that infant rhesus macaques, taught to share attention with a human through eye gaze or manual pointing, and then readily imitated such actions as tongue protrusion, clapping hands, hand clenching, and touching their own ears. It is also possible that mirror neurons become more finely tuned in the course of development.

In the monkey, mirror neurons fire when the monkey observes another individual reaching for an object, but not when the individual makes the same movement with no object present. That is, mirror neurons respond to transitive acts, but not to intransitive ones. In humans, in contrast, the mirror system, at least as understood through brain-imaging studies, appears to respond to both transitive and intransitive acts, perhaps paving the way to the understanding of acts that are symbolic rather than object-related (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). Even in the monkey,

E-mail address: [m.corballis@auckland.ac.nz](mailto:m.corballis@auckland.ac.nz)

<sup>1</sup> At the time of writing, a Google search yields just under 250,000 entries for mirror neurons, and around 39,000 for mirror neurons in conjunction with DNA.

though, the activity of mirror neurons is not restricted to exact physical mapping, and can even cross modalities. Some neurons in area F5, for example, discharge to the *sounds* of certain actions, such as the tearing of paper or the cracking of nuts (Kohler et al., 2002). This suggests that the mirror system comprises more than simple somatotopic mapping, and probably involves associative learning.

It is now recognized that mirror neurons do not function in isolation, but are rather part of a more extensive network involving other structures, including the superior temporal sulcus (STS) and inferior parietal lobule (area PF) (Rizzolatti et al., 2001). Homologs of these areas in the human brain appear to play a role in the imitation of action. Brain-imaging studies suggest, for example, that the superior temporal cortex, homolog of STS, provides a higher-order description of observed action (Puce & Perrett, 2003), while the rostral part of the posterior parietal cortex, homolog of PF, provides somatosensory information associated with the observed and to-be-executed action (Decety et al., 1997).

The discovery of the mirror system has also supported a long-held view that thought is “embodied” (e.g., Gibbs, 2006; Rizzolatti & Sinigaglia, 2008) or “grounded” (Barsalou, 2008), a view that contrasts with the more traditional view of the brain as an input–output device, with a central component representing thought. Embodied cognition implies instead that perception and movement are computationally intertwined, and that thought, rather than depending on amodal or abstract symbols, is construed as made up of simulations of real-world events, bodily states, and internal representations of actions. A succinct description is provided by the developmental psychologist Esther Thelen and colleagues:

To say that cognition is embodied means that it arises from bodily interactions with the world. From this point of view, cognition depends on the kinds of experiences that come from having a body with particular perceptual and motor capacities that are inseparably linked and that together form the matrix within which memory, emotion, language, and all other aspects of life are meshed. The contemporary notion of embodied cognition stands in contrast to the prevailing cognitivist stance which sees the mind as a device to manipulate symbols and is thus concerned with the formal rules and processes by which the symbols appropriately represent the world (Thelen, Schöner, Scheier, & Smith, 2001, p. 1).

Of course, not all of cognition is transparently embodied, and over time may develop a more symbolic character, in part through the process of conventionalization, as discussed below.

## 2. The case of language

These two extreme views of cognition pervade theories of language, which run from one extreme to the other. In early Chomskyan linguistics (e.g., Chomsky, 1957, 1975) and its more recent descendants (e.g., Nowak, Komorova, & Niyogi, 2002; Nowak, Plotkin, & Jansen, 2000), syntax is viewed as a computational system, involving manipulation of symbols, with little or no regard to their meaning. In contrast, what has come to be known as “cognitive linguistics” sees language as grounded in experience (e.g., Fauconnier, 1985; Langacker, 1987, 1991; Tomasello, 2003a)—this even includes abstract concepts assumed to be derived from experience but expressed in more bodily fashion through metaphor (Lakoff & Johnson, 1980). Love, for example, might be conveyed in terms of eating, as in “being consumed by love”, and argument in terms of war, as in “I demolished his argument”, or “The points he made were right on target”.

The mirror system provides a natural substrate for the embodiment of language. Mirror neurons were themselves first discovered in area F5 of the monkey’s frontal cortex, an area considered the homolog of Broca’s area (Rizzolatti & Arbib, 1998), which is the first area to have been identified as involved in language (Broca, 1861). More precisely, Broca’s area in humans can be divided into Brodmann areas 44 and 45, with area 44 considered the true analog of area F5. Brain-imaging shows that area 45 is activated by language output, whether spoken or signed (Horwitz et al., 2003), whereas area 44 is activated by nonlinguistic motor functions including complex hand movements, and sensorimotor learning and integration (Binkofski & Buccino, 2004). Indeed it has been proposed that “Broca’s area” should now be regarded as a collective term, involving many different functions, and no clearly demarcated subdivisions (Lindenberg, Fangerau, & Seitz, 2007).

The overlap between action imitation and language in the human brain appears to go beyond Broca’s area to the extended mirror system. Iacoboni and Wilson (2006) write that “It is quite striking that the three key brain areas identified for imitation are all considered crucial for language in classical neurolinguistic models (p. 504)”. Wernicke’s area, they suggest, performs a function in language analogous to that of the STS in imitation, and lesions to the rostral part of the posterior parietal cortex often result in conduction aphasia (Green & Howes, 1978). Such observations suggest that language is at least in part critically dependent on the mirror system itself (but see Grodzinsky, 2006, for some caveats). Evidence from fMRI shows that the mirror neuron region of the premotor cortex is activated not only when observers watch movements of the foot, hand, and mouth, but also when they read phrases pertaining to these movements (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006), despite the fact that words themselves do not map somatotopically onto the movements.

A striking feature of the brain areas involved in language, including Broca’s and Wernicke’s areas, is that they are represented predominantly in the left cerebral hemisphere. This is especially true of language production; comprehension is more bilaterally represented, at least at the word level and below (Hickok & Poeppel, 2000, 2004, 2007). One possibility is that a left-sided bias was introduced by the incorporation of vocalization. Although the primate mirror system is activated by the sounds of actions, it seems not to be activated by primate vocalizations themselves (Kohler et al., 2002), and there is no evidence that vocal production in nonhuman primates involves the mirror system, or indeed the motor cortex. Vocalization itself, though, appears to be left-hemispheric even in nonhuman species, from frogs (Bauer, 1993) to mice (Ehert, 1987) to primates (e.g., Hook-Costigan & Rogers, 1998), leading Corballis (2003) to propose that it was the incorporation of vocalization into the human mirror system that extended the left-hemispheric bias to the cortical level, giving rise perhaps to nonlinguistic asymmetries such as the left-hemispheric dominance in manual praxis—and indeed the phenomenon of right-handedness itself. Alternatively, it may have been simply the added complexity of language, and perhaps manual skill that resulted in the system becoming lateralized in humans.

In humans, there is some evidence that the mirror system is lateralized to the left hemisphere even for manual operations unrelated to language. Sekiyama, Miyauchi, Imaruoka, Egusa, and Tashiro (2000) report activation in Broca’s area when participants were given the task of identifying pictures of a rotated hand as either a left or a right hand, a task that evidently required them to imagine their own hand rotated to match that of the pictured hand. Activation in the inferior prefrontal cortex was restricted to the left hemisphere regardless of whether the depicted hand was a left or a right hand. Participants were also adapted to left–right reversing prisms, and were later able to use either the normal

or the reversed representation, but this too did not alter the left-hemispheric bias in activation. Aziz-Zadeh and colleagues report left-hemispheric activation in the human motor cortex when people listen to the sounds of manual actions (Aziz-Zadeh, Iacoboni, Zaidel, Wilson, & Mazziotta, 2004) but Broca's area and its right-hemispheric homolog were equally activated during imitation of finger movements (Aziz-Zadeh, Koski, Zaidel, Mazziotta, & Iacoboni, 2006). Yet Fecteau, Lassonde, and Théoret (2005) report that only the left hemisphere was activated in a split-brained patient by action observation, suggesting that right-hemisphere activation in intact participants may be a result of callosal transfer from the left hemisphere.

In summary, the mirror system in humans includes a number of properties not evident in nonhuman primates. It appears to be involved in several aspects of language, deriving initially perhaps from its capacity to respond to intransitive as well as transitive actions. Language functions related to the mirror system are predominantly left-hemispheric, at least in the majority of individuals. This bias may carry over into at least some nonlinguistic manual activities.

### 3. The gestural theory of language evolution

At some point in hominin evolution, then, the mirror system has assumed an additional role in language and become lateralized, while remaining rooted in manual activity. Moreover language itself need not involve vocalization—the signed languages of the deaf are now known to have all of the linguistic sophistication of spoken languages (e.g., Emmorey, 2002; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000). As noted above, Broca's area is activated by both speaking and signing (Horwitz et al., 2003). These observations are consistent with the speculation that language evolved from manual gestures rather than from vocal calls (e.g., Arbib, 2005a, 2005b; Armstrong & Wilcox, 2007; Corballis, 2002, 2004b; Fogassi & Ferrari, 2007; Rizzolatti & Arbib, 1998; Rizzolatti & Sinigaglia, 2008).<sup>2</sup>

Further, as Rizzolatti and Sinigaglia point out, the neural circuits involved in the generation of vocal calls in nonhuman primates are radically different from those involved in human speech, and involve the cingulate cortex, diencephalon, and brainstem structures, rather than the classical cortical speech areas (Jürgens, 2002). Although vocal communication is widespread in the animal kingdom, vocalizations are for the most part genetically structured and intentional control is limited, at best. Surprisingly few species seem capable of vocal learning, which is of course critical to spoken language. These species include elephants, seals, killer whales, and some birds, but of the primates only humans are vocal learners (Jarvis, 2006). In marked contrast to the inflexibility of vocalizations in primates, the communicative bodily gestures of gorillas (Pika, Liebal, & Tomasello, 2003), chimpanzees (Liebal, Call, & Tomasello, 2004), and bonobos (Pika, Liebal, & Tomasello, 2005) are subject to social learning and sensitive to the attentional state of the recipient—both prerequisites for language. In chimpanzees and bonobos, bodily gestures are much less tied to context—another prerequisite for language—than are vocalizations (Pollick & de Waal, 2007). Our great-ape heritage therefore ensured that our hominin forebears were much better pre-adapted to intentional communication based on manual and other bodily gestures than to one based on vocalization.

This conclusion is reinforced by attempts to teach language to great apes. Early efforts made it clear that apes are unable to produce anything approaching human speech. In one famous example, a baby chimpanzee reared in a human family proved able to articulate only three or four words, and these were actually whispered rather than vocalized. His verbal skills were soon eclipsed by those of the human children in the family (Hayes, 1952). It was then realized that the failure to speak may have resulted from deficiencies of the vocal apparatus, and perhaps of cortical control of vocal output, rather than a failure of language itself. Subsequent attempts to teach language to nonhuman primates have therefore been based on manual action and visual representations. For example, the chimpanzee Washoe was taught over 100 manual signs,<sup>3</sup> based loosely on American Sign Language (ASL), and was able to combine signs into two- or three-“word” sequences to make simple requests (Gardner & Gardner, 1969). Kanzi, a bonobo, has an even larger vocabulary, based partly on pointing to symbols on a keyboard and supplemented with manual gestures, but his ability to construct meaningful sequences appears to be limited to only two or three “words”.

Kanzi has nevertheless shown an ability to follow instructions conveyed in spoken sentences with as many as seven or eight words (Savage-Rumbaugh, Shanker, & Taylor, 1998). Even more impressively, a border collie known as Rico has been shown to follow spoken instructions to fetch a particular item from a selection of objects and either place it in a box or bring it to a particular person (Kaminsky, Call, & Fischer, 2004). When given a name he has not previously heard, he then selects the one novel object placed among the alternatives, and thenceforth is able to apply that label to that object. This rapid association of arbitrary labels to objects, generally considered unique to human infants, is known as “fast mapping”. Besides suggesting that the association of spoken words with objects or actions may not be exclusive to humans, these exploits may cast some doubt on the notion that the mirror system is involved in language comprehension. After all, neither Kanzi nor Rico can speak, so they are presumably without any mirror neurons that map spoken words onto their production.

Of course, stimuli can act as cues for action without any involvement of the mirror system, as when a dog obeys a whistle when rounding sheep. Kanzi's accuracy in following instructions was shown to be roughly equivalent to that of a 2½-year-old girl (Savage-Rumbaugh et al., 1998), and is probably based on the extraction of two or three key words rather than a full decoding of the syntax of the sentences. These key words may act as discriminative stimuli to trigger behavior, and presumably do not imply “action understanding” in the sense that Rizzolatti et al. (2001) used the term, or even in the sense of mirror neurons responding to the sounds of actions (Kohler et al., 2002). Bloom (2004) has similarly suggested that Rico may have been using the spoken words as cues for fetching routines, without any broader sense of reference. Further research may provide more specific information as to what these animals extract from spoken words, but it is at least clear that their exploits do not constitute true language, which involves both production and perception.

### 4. From mimesis to conventionalization

Our hominin forebears, then, were much better adapted to communicate intentionally using bodily movements, and especially movements of the hands, than to vocalize intentionally. The trainers of Kanzi and other great apes chose deliberately to use noniconic

<sup>2</sup> Indeed, there was fairly compelling evidence for the manual origins of language even before mirror neurons were discovered, or became widely known (e.g., Armstrong, Stokoe, & Wilcox, 1995; Condillac, 1746/1971; Corballis, 1991; Hewes, 1973; Wundt, 1921). Of course, there are some who still disagree (e.g., Burling, 2005; Cheney & Seyfarth, 2005; MacNeilage, 2008).

<sup>3</sup> Washoe did not easily acquire these signs, which had to be molded by her teachers. This is further evidence that adult nonhuman primates do not readily imitate.

systems of representation, in the hope of demonstrating something approximating human language, but it seems reasonable to suppose that the precursors of human language were based on mimesis. The body and hands are free to move in four dimensions (three of space and one of time), with many more degrees of freedom, and so mimic activity in the external world. Donald (1991) suggests that what he calls “mimetic culture” may have evolved with *Homo erectus* from some 2 million years ago; it may have depended in large part on obligate bipedalism, which freed the hands from involvement in locomotion. The bipedalism of the earlier Australopithecines is considered facultative rather than obligate, and perhaps had a lesser influence. The hands can also assume, at least approximately, the shapes of objects or animals, and the fingers can mimic the movement of legs and arms. The movements of the hands can also mimic the movement of objects through space, and facial expressions can convey something of the emotions of events being described. Mimesis persists in dance, ballet, and mime, and we may all resort to mime when trying to communicate with people who speak a language different from our own. Once, in Russia, I was able to successfully request a bottle opener by miming the action of opening a beer bottle, to the vast amusement of the people at the hotel desk.

The modern sign languages of the deaf are also partially dependent on mime. It has been estimated, for example, that in Italian Sign Language some 50% of the hand signs and 67% of the bodily locations of signs stem from iconic representations, in which there is a degree of spatiotemporal mapping between the sign and its meaning (Pietrandrea, 2002). Emmorey (2002) notes that in ASL some signs are arbitrary, but many more are iconic. For example, the sign for “erase” resembles the action of erasing a blackboard, and the sign for “play piano” mimics the action of actually playing a piano. But of course signs need not be transparently iconic, and the meanings of even iconic symbols often cannot be guessed by naïve observers (Pizutto & Volterra, 2000). Moreover, iconic signs are not treated simply as iconic gestures, and signers often do not even recognize their iconic character. Signs also tend to become less iconic and more arbitrary over historical time, in the interests of speed, efficiency, and grammatical constraints. This process is known as *conventionalization* (Burling, 1999).<sup>4</sup>

The Swiss linguist Ferdinand de Saussure (1916/1977) wrote of the “arbitrariness of the sign” as a defining property of language, and on this basis it is sometimes supposed that signed languages, with their strong basis in iconic representations, are not true languages. Although most words in spoken languages are indeed arbitrary—the words *cat* and *dog* in no way resemble those animals or the sounds that they make—there are of course some words that are onomatopoeic. One such word is *zanzara*, which is the evocative Italian word for mosquito, and Pinker (2007) notes a number of newly minted examples: *oink*, *tinkle*, *barf*, *conk*, *woofer*, *tweeter*. The arbitrariness of words (or morphemes) is not so much a necessary property of language, though, as a matter of expedience, and of the constraints imposed by the language medium. Speech, for example, requires that the information be *linearized*, squeezed into a sequence of sounds that are necessarily limited in terms of how they can capture the physical nature of what they represent. The linguist Charles Hockett (1978) put it this way:

... when a representation of some four-dimensional hunk of life has to be compressed into the single dimension of speech, most iconicity is necessarily squeezed out. In one-dimensional projections, an elephant is indistinguishable from a woodshed. Speech perforce is largely arbitrary; if we speakers take pride

in that, it is because in 50,000 years or so of talking we have learned to make a virtue of necessity (pp. 274–275).

The symbols of signed languages are clearly less constrained. Since the hands and arms can, to some extent at least, mimic the shapes of real-world objects and actions, lexical information can be delivered in parallel instead of being forced into rigid temporal sequence. Even so, conventionalization allows signs to be simplified and speeded up, to the point that many of them lose most or all of their iconic aspect. For example, the ASL sign for *home* was once a combination of the sign for *eat*, which is a bunched hand touching the mouth, and the sign for *sleep*, which is a flat hand on the cheek. Now it consists of two quick touches on the cheek, both with a bunched handshape, so the original iconic components are effectively lost (Frishberg, 1975).

Along with conventionalization, the role of the mirror system itself may have shifted. With pure mimesis, the action understanding arising from mimed action is more or less directly mapped onto the action itself. Thus a skilled mime artist may be able to simulate riding a bicycle in the absence of the bicycle itself, and the mime is immediately understood. But with actions that are more symbolic, the mapping depends on learned associations. The receiver understands the action of the sender in terms of how she, the viewer, would perform the action, and this understanding is then associated with the meaning of the action. Even then, though, the mapping may be grounded in movement. As we have seen, spoken phrases about movements of the foot, hand, activate the corresponding mirror-neuron regions of the premotor cortex (Aziz-Zadeh et al., 2006).

## 5. From hand to mouth

Once language is conventionalized and loses its mimetic aspect, then, it need not be restricted to the visual modality. In the course of human evolution, vocalizations have replaced manual acts as the main medium of language, although manual languages persist, especially in the form of signed languages invented by deaf communities. The indigenous Australians also developed signed languages, although these are based on spoken languages and are used to overcome speech taboos associated with certain ceremonies (Kendon, 1988). A form of signed language known as Plains Sign Talk was developed by native Americans, probably as a *lingua franca* to allow different tribes to communicate with each other, and one dictionary lists over 3000 signs (Mallery, 1880). Signing may also have been prominent among certain African tribes, as suggested by the following provocative quote from the 19th-century British explorer, Mary Kingsley (1897/1965):

[African languages are not elaborate enough] to enable a native to state his exact thought. Some of them are very dependent upon gesture. When I was with the Fans they frequently said “We will go to the fire so that we can see what they say”, when any question had to be decided after dark, and the inhabitants of Fernando Po, the Bubis, are quite unable to converse with each other unless they have sufficient light to see the accompanying gestures of the conversation (p. 504).<sup>5</sup>

Of course, even in modern societies, people with normal hearing continue to gesture with their hands while they speak (McNeill, 1992; Goldin-Meadow & McNeill, 1999; Willems & Hagoort, 2007). Indeed, manual gesture may lurk not far below the surface;

<sup>5</sup> This is no doubt a highly Eurocentric account, and even if Kingsley's observations were accurate there is no reason to suppose the gestural (or combined vocal/manual) languages used by these tribes were in any sense linguistically inferior. Kingsley wrote at a time when signed languages were widely regarded as inferior to speech.

<sup>4</sup> Conventionalization is a historical process, not a developmental one. Children readily learn abstract signs or words without first having to encounter iconic versions.

when asked to communicate without speaking, hearing adults spontaneously develop a form of sign language, in which grammatical components are introduced (Goldin-Meadow, McNeill, & Singleton, 1996). Manual gestures also play an important role in the development of normal speech (e.g., Bates & Dick, 2002). For example, canonical babbling in children aged from 6 to 8 months is accompanied by rhythmic hand movements (Masataka, 2001). Word comprehension in children between 8 and 10 months and word productions between 11 and 13 months are accompanied by gestures of pointing and showing and gestures indicating recognition, respectively (Bates & Snyder, 1987). Manual gestures predate early development of speech in children, and predict later success even up to the two-word level (Iverson & Goldin-Meadow, 2005).

Despite the continuing presence of a manual component, the notion that language switched from manual to vocal remains somewhat speculative and to some seems too large a transition to be plausible. The linguist Robbins Burling (2005), for example writes

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

In another recent book, MacNeilage (2008) expresses similar concerns, but I will argue that the switch was probably a relatively simple and natural one.

First, as we have seen, mirror neurons can be activated by acoustic stimuli, such as the tearing of paper or the cracking of nuts (Kohler et al., 2002). The important point here is that, even in the monkey, the mirror system is grounded in action, not in the particular modality of input. Moreover, speech itself can be viewed as a gestural system rather than an acoustic one (Studdert-Kennedy, 1998). As Rizzolatti and Arbib (1998) recognized, the mirror system operates according to the same principles as postulated earlier by the motor theory of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), which holds that speech sounds are perceived in terms of how they are produced, rather than as acoustic elements. As Galantucci, Fowler, and Turvey (2006) put it in a recent review of the motor theory, “perceiving speech is perceiving gestures (p. 361).”<sup>6</sup> This has led to what is known as articulatory phonology (Browman & Goldstein, 1995), in which speech is understood as gestures produced by six articulatory organs, the lips, the velum, the larynx, and the blade, body, and root of the tongue (see also Goldstein, Byrd, & Saltzman, 2006). In this context, of course, the term “gesture” is not confined to movements that are visible.

The motor theory of speech perception does not imply that listening to speech induces speech gestures themselves, just as the firing of mirror neurons in the monkey does not cause the animal to actually perform an observed gesture. Updating the motor theory in 1985, Liberman and Mattingly (1985) state that the objects of speech perception are not the actual speech gestures, but rather the *intended* gestures. This might explain why people with Broca’s aphasia typically retain reasonably accurate perception and of speech, despite being unable to produce coherent speech. (An alternative account, though, is to reverse the motor theory of speech perception, and suppose that speech production is driven by auditory targets—see footnote 6). Nevertheless, some patients with frontal lesions do show marked deficits in both the production and comprehension of phonemes (Blumstein, Baker, & Good-

glass, 1977), and fMRI studies have shown that areas active when people listen passively to speech sounds overlap with those active in speech production (Pulvermüller et al., 2006; Wilson, Saygin, Sereno, & Iacoboni, 2004).

Thus speech and manual gestures appear to share a grounding in the mirror system, whereby perception is understood in terms of intentional action. But the links between the two may be even closer. In humans, grasping movements of the hand affect the kinematics of speech itself, in a fashion suggestive of the involvement of mirror neurons. Thus varying the size of objects that are grasped or brought to the mouth induces changes in parameters of lip kinematics and voice spectra of syllables pronounced simultaneously with the actions in both the person performing these actions, and in someone observing the same actions; the larger the object grasped, the wider the opening of the mouth, with consequent effects on speech formants (see Gentilucci & Corballis, 2006, for review). Subsequent work has revealed similar effects in 11- to 13-month-old infants (Bernardis, Bello, Pettenati, Stefanini, & Gentilucci, 2008).

This is not to say that speech and hand movements cannot be dissociated. Watkins, Strafella, and Paus (2003) report evidence from transcortical magnetic stimulation (TMS) that the motor system involved in speech production is activated both by listening to speech and watching speech movements, but there was no effect on motor-evoked potentials in the hand muscles. But facial gestures themselves may be critical to the switch between manual gestures and speech, and provide a bridge from one to the other. Manual and facial gestures are closely linked neurophysiologically, as well as behaviorally. For example, Rizzolatti et al. (1988) recorded from neurons in area F5 in the monkey that fire when the animal makes movements to grasp an object with either the hand or the mouth. Petrides, Cadoret, and Mackey (2005) have identified an area in the monkey brain just rostral to premotor area 6, also considered a homolog of part of Broca’s area that is involved in control of the orofacial musculature. Gentilucci, Benuzzi, Gangitano, and Grimaldi (2001) showed that when people were instructed to open their mouths while grasping objects, the size of the mouth opening increased with the size of the grasped object, and conversely, when they opened their hands while grasping objects with their mouths, the size of the hand opening also increased with the size of the object. These neural links between hand and mouth may derive from ingestive behavior rather than communication, perhaps involved in preparing the mouth to grasp an object after the hand has grasped it, but later adapted for gestural and finally vocal language.

Facial gestures themselves, perceived as visual stimuli rather than as speech, play a role in speech perception. Using fMRI, Calvert and Campbell (2003) have shown that watching speech activates cortical speech areas, including Broca’s area, the left superior temporal sulcus (STS), and the left supramarginal gyrus (which is part of Wernicke’s area). Deaf people often become highly proficient at lip-reading, and the visual component of speech is also illustrated by the McGurk effect (McGurk & MacDonald, 1976), in which dubbing sounds onto a mouth that is saying something different alters what the hearer actually hears; that is, the viewer/listener often reports what the speaker is seen to be saying rather than the speech sound itself, or sometimes a blend of the two. Of course we can adequately perceive speech when visual cues are absent, as in listening to radio or a telephone message, but there is little doubt that facial cues, when present, exert a significant influence.

It is also increasingly recognized that signed languages involve movements of the face as well as of the hands. Facial expressions and head movements can turn an affirmative sentence into a negation, or a question. Mouth gestures are especially important, to the point that some linguists are beginning to identify a form of

<sup>6</sup> It should be noted that the motor theory of speech perception remains controversial, especially among speech scientists, as Galantucci et al. (2006) recognize—see Hickok (in press) and Lotto, Hickok, and Holt (2009). An alternative, suggested to me by Gregory Hickok (see also Guenther, 2006) is that speech perception drives production, rather than the other way about. This implies a perceptual theory of speech production rather than a motor theory of speech perception!

phonology underlying mouth movements. Explicit schemes for mouthed “phonemes” have been proposed for a number of European Sign Languages, including Swedish, English, and Italian (Sutton-Spence & Boyes-Braem, 2001). Mouth gestures can also serve to disambiguate hand gestures, and as part of more general facial gestures provide the visual equivalent of prosody in speech (Emmorey, 2002). This work is still in its infancy, but suggests an evolutionary scenario in which mouth movements gradually assumed dominance over hand movements, and were eventually accompanied by voicing and movements of the tongue and vocal tract. That is, the progression was from hand to face to voice, but presumably with overlap at all stages. To this day, speech is characteristically accompanied by manual gestures, and can be partially “read” by observing facial movements.

The transitional role of the face may also help explain how vocalization was eventually introduced into the system. The face may be taken to include the tongue and other areas internal to the mouth, so the gradual addition of sound would make certain facial gestures more accessible, and indeed distinctive. It provides, for example, for the distinction between voiced and unvoiced consonants. Nevertheless, given that vocalization does not appear to be represented in the mirror systems of nonhuman primates, its incorporation was probably a gradual process of natural selection, involving a number of neurophysiological and anatomical changes.

## 6. When did speech predominate?

The “sticking point,” then, to borrow Burling’s (2005) phrase, was not so much the switch from visual to audible language as the incorporation of vocalization into the mirror system. In the monkey, a small number of neurons in lateral F5 do appear to respond to both the action and observation of mouth gestures, such as lipsmacking or protrusion of the lips and tongue, and these neurons have been dubbed “communicative mirror neurons” (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003). But these actions are primarily visual, and do not involve vocalization; as Ferrari et al. note “Communicative mirror neurons could be an evolutionary precursor of social communication mediated by facial gestures (pp. 660–661)”. In humans, in contrast, Hickok and colleagues have shown that both the production and perception of vocal speech activate both frontal and temporo-parietal areas, which might be taken to mean that vocal speech is incorporated into the mirror system, predominantly in the left hemisphere (Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Hickok & Poeppel, 2000; Vigneau et al., 2006)—although Hickok and colleagues do not interpret their findings in this way.

Precisely when vocal control was introduced into the mirror system is not clear. One clue comes from genetics. A mutation of the forkhead box transcription factor, FOXP2, in some members of an English family known as the KE family has resulted in a quite severe deficit in vocal articulation (Watkins, Dronkers, & Vargha-Khadem, 2002). Moreover, the members of the KE family affected by the mutation, unlike their unaffected relatives, show no activation in Broca’s area while covertly generating verbs (Liégeois et al., 2003). This might be taken to mean that the FOXP2 gene in humans is involved in the cooption of vocal control by Broca’s area (Corballis, 2004a). In songbirds, knockdown of the FOXP2 gene impairs the imitation of song (Haesler et al., 2007), and insertion of the FOXP2 point mutation found in the KE family into the mouse critically impairs synaptic plasticity and motor learning (Groszer et al., 2008). Its role in foxes is unknown.

Although highly conserved in mammals, the FOXP2 gene underwent two mutations since the split between hominid and chimpanzee lines. According to one theoretical estimate, the more recent of these occurred “some 10,000–100,000 years ago” (Enard et al., 2002), although the error associated with this estimate

makes it not unreasonable to suppose that it coincided with the emergence of *Homo sapiens* around 170,000 years ago. Enard et al. (2002) write that their dating of the FOXP2 mutation “is compatible with a model in which the expansion of modern humans was driven by the appearance of a more-proficient spoken language” (p. 871).

This conclusion is challenged, though, by recent evidence that the mutation is also present in the DNA of a 45,000-year-old Neandertal fossil, suggesting that it goes back at least 300,000–400,000 years to the common ancestor of humans and Neandertals (Krause et al., 2007). But this is challenged in turn by Coop, Bullaughev, Luca, and Przeworski (2008), who used phylogenetic dating of the haplotype to re-estimate the time of the most recent common ancestor carrying the FOXP2 mutation; their answer of 42,000 years ago is much more consistent with the estimate reported by Enard et al. (2002) than with the estimate implied by Krause et al. (2007). They argue that the presence of the mutation in Neandertal was more likely due to contamination of the Neandertal DNA, or to low rates of gene flow between human and Neandertal, on the assumption that the allele was globally beneficial, or. This suggestion is not without precedent; other evidence suggests that microcephalin, a gene involved in regulating brain size, may have entered the human gene pool through interbreeding with Neandertals (Evans, Mekel-Bobrov, Yallender, Hudson, & Lahn, 2006), so the reverse possibility of FOXP2 entering the late Neandertal gene pool from *H. sapiens* is not completely ruled out. We might have been slightly more chummy with the Neandertals than is generally thought.

There is much yet to be clarified about the FOXP story. Although at least some of the evidence reviewed above suggests that FOXP2 is involved in vocalization, its precise role is unclear. The mutation in the KE family, for example, is unrelated related to that which occurred in hominin evolution, and FOXP2 is implicated in many parts of the body, such as the gut, and in regions of the brain not implicated in the mirror system, such as the basal ganglia. In any event, other factors were surely involved in the switch from manual gesture to speech. Fossil evidence suggests that the anatomical requirements for fully articulate speech were probably not complete until the emergence of *H. sapiens*. For example, the hypoglossal nerve, which passes through this canal and innervates the tongue, passes through the hypoglossal canal, and this canal is much larger in humans than in great apes, probably because of the important role of the tongue in speech. Fossil evidence suggests that the size of the hypoglossal canal in early australopithecines, and perhaps in *Homo habilis*, was within the range of that in modern great apes, while that of the Neandertal and early *H. sapiens* skulls was contained well within the modern human range (Kay, Cartmill, & Barlow, 1998), although this has been disputed (DeGusta, Gilbert, & Turner, 1999). A further clue comes from the finding that the thoracic region of the spinal cord is relatively larger in humans than in nonhuman primates, probably because breathing during speech involves extra muscles of the thorax and abdomen. Fossil evidence indicates that this enlargement was not present in the early hominids or even in *Homo ergaster*, dating from about 1.6 million years ago, but was present in several Neandertal fossils (MacLarnon & Hewitt, 2004).

According to P. Lieberman (1998; Lieberman, Crelin, & Klatt, 1972), however, the lowering of the larynx necessary for the full range of speech sounds was incomplete in the Neandertals.<sup>7</sup> This

<sup>7</sup> Robert McCarthy of Florida Atlantic University has recently simulated how the Neandertal would have sounded when articulating the syllable /ij/ (or ee), based on the shape of the vocal tract. It can be found on <http://www.anthropology.net/2008/04/16/reconstructing-neandertal-vocalizations/>, and compared with a human articulating the same sound. One observer described the Neandertal’s attempt as sounding more like a sheep or a goat than a human.

work has been controversial (e.g., Boë, Heim, Honda, & Maeda, 2002; Gibson & Jessee, 1999), but there is other evidence that the cranial structure underwent changes subsequent to the split between anatomically modern and earlier “archaic” *Homo*, such as the Neanderthals, *Homo heidelbergensis*, and *Homo rhodesiensis*. One such change is the shortening of the sphenoid, the central bone of the cranial base from which the face grows forward, resulting in a flattened face (D. E. Lieberman, 1998). Lieberman, McBratney, and Krovitz (2002) speculate that this is an adaptation for speech, contributing to the unique proportions of the human vocal tract, in which the horizontal and vertical components are roughly equal in length. This configuration, he argues, improves the ability to produce acoustically distinct speech sounds, such as the vowel [i]. It is not seen in Neanderthal skeletal structure (see also Vleck, 1970). Another adaptation unique to *H. sapiens* is neurocranial globularity, defined as the roundness of the cranial vault in the sagittal, coronal, and transverse planes, which is likely to have increased the relative size of the temporal and/or frontal lobes relative to other parts of the brain (Lieberman et al., 2002). These changes may reflect more refined control of articulation and also, perhaps, more accurate perceptual discrimination of articulated sounds.

The most persistent advocate of the late emergence of speech is Lieberman, and as recently as 2007 he summarized as follows:

... fully human speech anatomy first appears in the fossil record in the Upper Paleolithic (about 50,000 years ago) and is absent in both Neanderthals and earlier humans (Lieberman, 2007, p. 39).

This provocative statement suggests that articulate speech emerged even later than the arrival of *H. sapiens* some 200,000 years ago. While this may be an extreme conclusion, the bulk of evidence does suggest that autonomous speech did emerge very late in the human repertoire.

In summary, the key to the evolution of speech may have been the incorporation of vocalization into the mirror system, and thus the capacity to use vocalization as part of the intentional system. The evidence suggests that this was complete late in hominin evolution, and possibly only in our own species *H. sapiens*. It is likely, though, that this was not a sudden process.

## 7. A gradual switch?

Some authors, including P. Lieberman (e.g., 1998) himself, have proposed that language itself evolved suddenly, presumably on the assumption that language is equated with speech. This is sometimes termed the “big bang” theory of language evolution, and attributed to Bickerton (1995), who once wrote that “... true language, via the emergence of syntax, was a catastrophic event, occurring within the first few generations of *H. sapiens sapiens* (p. 69)”. But language is highly complex, which makes it unlikely that it evolved in a single saltational step. The assumption that it evolved from manual and facial gestures allows us to consider a more gradual and evolutionarily realistic progression, going back perhaps 2 million years to the origins of the genus *Homo*.

Arbib (2005a and 2005b) has suggested that gestural communication did not progress to full language before the introduction of vocalization. Rather, gestural communication progressed to the level of what he called “protosign”, which is the stage at which the mirror system has moved from the understanding of transitive action to the intransitive use of action for communication. Effectively, this is also the mimetic stage proposed by Donald (1991), or the stage reached by Kanzi and Washoe. Arbib suggests that protosign served as a scaffold for protospeech, and protosign and protospeech then co-evolved in an “expanding spiral” toward true language. This scenario is similar to the one envisaged here, and

perhaps indistinguishable from it. My own view, though, is that the switch from protolanguage to language was itself gradual—and even diffuse, as I explain below—and occurred in parallel with the switch from a primarily manual form of communication to a primarily vocal one, with varying degrees of admixture. The distinction between protolanguage and language is generally depicted as all-or-none, but this has recently been challenged, for example by Smith (2008) who argues that protolanguage probably evolved into complex language gradually, through the same processes that underlie on-going changes in present-day languages. The concept of a saltational switch from protolanguage to language was perhaps encouraged by the notion that grammatical language depends on an innate, uniquely human endowment known as universal grammar, or UG (e.g., Chomsky, 1975), but UG itself has come under increasing attack, on logical, evolutionary, and empirical grounds (e.g., Christiansen & Chater, 2008; Everett, 2005, 2007; Tomasello, 2003a; Wray, 2002). Contrary to the concept of UG, evidence from the world’s languages, including those of nonliterate cultures, suggests that there are few if any grammatical constructions that are universally present in all languages (Tomasello, 2003b).

Christiansen and Chater (2008), following Deacon (1997), suggest that, rather than shaping the brain, language was shaped by the brain, and is therefore dependent on the nature of human thought processes, cognitive limitations, cultural influences and the like. This is compatible with the notion, discussed earlier, that cognition in general and language in particular, is based more on human action and experience than on an abstract symbol-processing system. More importantly for the present discussion, grammar is seen to adapt over time, in a process known as *grammaticalization* (Heine & Kuteva, 2007; Hopper & Traugott, 1993). The shaping of language to different cultural necessities, and different modalities of expression, may therefore have been a gradual, flexible, and incremental process.

Grammaticalization involves the gradual shift from content words, expressing meaning, to function words, serving purely grammatical functions. According to the 18th-century English philologist John Horne Tooke (1857), the earliest “language” consisted only of nouns and verbs (“necessary words”), while other word classes, such as adjectives, adverbs, prepositions, and conjunctions arose from the abbreviation or “mutilation” of these necessary words. Adjectives, for example, may derive from nouns, as in *heavenly* or *manly*, where the suffix *-ly* is presumably a contraction of *-like*. This idea is endorsed by Hurford (2003), who gives an example from the emergence of Tok Pisin, a creole that derived from pidgin in Papua New Guinea. Pidgins are makeshift languages created as a means of communication between speakers of different languages and have little or no grammar. They often develop into creoles, which have more sophisticated grammars. The Papua New Guinean pidgin consisted only of nouns and verbs, but in Tok Pisin adjectives were signaled by the addition of the suffix *-fela* (or *-pela*), itself derived from the English noun *fellow*. It also seems reasonable to suppose that phrase structure preceded the ability to combine phrases into single utterances. Christiansen and Kirby (2003) give the example of the phrases *My dad/He plays tennis/He plays tennis with his colleagues*, which can be combined into the more compact form *My dad plays tennis with his colleagues*. These and other aspects of grammar are unlikely to have emerged in a single, fortuitous mutation, but were more likely the products of gradual, culturally-driven attempts to make language more efficient and economical.

A living example of the development of grammar comes from the emergence of a new sign language, called Lenguaje de Signos Nicaragüense (LSN), among the deaf community in Nicaragua. In the course of time, LSN has changed from a system of holistic signs to a more combinatorial format. For example, one generation of

children were told a story of a cat that swallowed a bowling ball, and then rolled down a steep street in a “waving, wobbling manner”. The children were then asked to sign the motion. Some indicated the motion holistically, moving the hand downward in a waving motion. Others, however, segmented the motion into two signs, one representing downward motion and the other representing the waving motion, and this version increased after the first cohort of children had moved through the school (Senghas, Kita, & Özyürek, 2004). Segmentation allows the components to be used in new combinations, thereby increasing the flexibility and efficiency of the system.

## 8. On time

We saw earlier that the brain areas involved in human language overlap quite extensively with the homolog of the extended mirror system as identified in monkeys (Iacoboni & Wilson, 2006), and indeed with the areas in the human brain involved in imitation. This suggests that differentiation within the mirror system may have provided for the evolution of language, and indeed suggests that the mirror system itself should not be regarded as a unitary entity, at least in the human brain. Yet there must surely be aspects of language and of its evolution that lie outside the mirror system.

One potentially important cognitive development that may have had a prominent, if not definitive, influence on the evolution of language was the capacity to travel mentally in time, as reflected in episodic memory and the capacity to imagine detailed future scenarios. Arguably uniquely human, mental time travel may well have evolved with the emergence of the genus *Homo* from some 2 million years ago, as an adaptation enabling more detailed recording of critical environmental events, and the ability to compare future scenarios for optimal selection of strategy (Suddendorf & Corballis, 1997, 2007). At least some of the features of grammatical language may derive from pressure to communicate events, past, present, or planned, so that their benefits can be shared, enhancing group survival in the dangerous environment of the savanna (Corballis & Suddendorf, 2007). In this spirit, Gärdenfors (2004) writes that “there has been a co-evolution of cooperation about future goals and symbolic communication” (p. 243).

Reference to events that are remote from the present requires an extensive system of storage and vocabulary. The average literate person is said to have around 50,000 concepts, along with words to refer to them (Pinker, 2007). This vocabulary was probably driven in part by the requirements of mental time travel, so we can conjure up and describe events that are distant in both time and space. Language also needs ways of expressing time itself, and different languages do this in different ways. In English and other European languages, aspects of time are coded in tense—adjustments to the verb to indicate whether events are located in the past, present, or future, as well as other temporal aspects such as whether an action is completed or on-going. Chinese, though, has no tenses, and the time of an event can be indicated by adverbs, such as *tomorrow*, or by what are called aspectual markers, as in a sentence that might be roughly rendered as *He break his leg before* (Lin, 2005). In ASL, time is represented as a line from the back to the front of the body, with the past at the back and the future in front. This is also implicit in English prepositions, where time is built on spatial metaphor; a sentence like *I arrived after him* means that I arrived later in time. In Chinese, time is represented vertically, and travels downwards, so that *the month above* means last month (Chen, 2007). The Pirahã, a tribe of some 200 people in Brazil, have a relatively primitive way of indicating time, in the form of two tense-like morphemes which seem to indicate simply whether an event is in the present or not; correspondingly, the Pirahã are said to live largely in the present (Everett, 2005). Revealingly, Everett writes:

... these apparently disjointed facts about the Pirahã language—gaps that are very surprising from just about any grammarian’s perspective—ultimately derive from a single cultural constraint in Pirahã, namely, *the restriction of communication to the immediate experience of the interlocutors* (p. 622; Everett’s italics)

Mental time travel is also generative, which may explain the generativity of human language itself. In planning for the future, we have the capacity to create and compare different scenarios, built from our storehouse of semantic concepts and episodic memories. Even episodic memory itself is better regarded as a construction, not always accurate; as Neisser (2008) put it, “Remembering is not like playing back a tape or looking at a picture; it is more like telling a story” (p. 88). And a pervasive product of our capacity for generativity is fiction itself. Humans seem to have a limitless thirst for stories, whether told, written, or enacted, or whether true or false, and language is the principal medium through which these are shared.

The brain areas involved in mental time travel appear to lie largely outside of those involved in language per se. Schacter, Addis, and Buckner (2008) refer to a “core network” that deals with the simulation of future events based on memory for past episodes, and this network is activated both in remembering the past and imagining the future. Buckner, Andrews-Hanna, and Schacter (2008) refer somewhat more generally to the “brain’s default network”, active when the mind cuts off from immediate sensory input and focuses inwards—reconstructing the past, envisaging the future, and imagining the perspectives of others. These largely overlapping networks include prefrontal and medial temporal regions, and posterior regions including posterior cingulate and retrosplenial cortex. They seem almost fastidious in avoiding the mirror system. Together, the mirror system and the default network may provide a broader context within which to understand the complementary functions of language and mind.

## 9. Conclusion

If language is indeed shaped by cognitive structures, such as the representation and understanding of time, and by cultural imperatives, then there may be no specific point at which “language” may be said to have emerged. The extent of cultural differences may be illustrated in part by the differences between the language of the Pirahã, mentioned above, and modern European languages. Pirahã language has no numbers or system of counting, no color terms, no perfect tense and, as we have seen, only a very primitive way of talking about relative time. It may even be said to lack verbs, in the sense of a verb as a linguistic class; the Pirahã learn verbs one by one as individual entities. There are no embedded clauses (Everett, 2005). One might be tempted to believe that they suffer from some genetic defect, but this idea is rejected by Everett, who describes them as “some of the brightest, pleasantest, and most fun-loving people that I know” (p. 621).<sup>8</sup>

The Pirahã language is nevertheless rich in morphology and prosody, and Everett insists that it should not be regarded as in any way “primitive”, and suggests that it is probably not especially unusual; other languages of nonliterate peoples may have similar characteristics. The point is that language adapts to serve the needs of the culture, and varies widely from one culture to another. In this variation, as Christiansen and Chater (2008) make clear, there is no space for UG in any nontrivial sense.

Once speech became the dominant mode, language may well have changed in significant ways, quite apart from the mode change itself. The emergence of autonomous speech would have

<sup>8</sup> A critique of Everett’s (2005) paper can be found in Nevins, Pesetsky, and Rodrigues (2007), and a reply in Everett (2007).



freed the hands from involvement in other activities, such as carrying things, or manufacturing tools or other artifacts. I have suggested elsewhere (Corballis, 2004b) that this may explain what has been termed the “human revolution” (Mellars & Stringer, 1989), apparently restricted to our own species, comprising sudden and rapid advances in technology, culture, art, bodily ornamentation, and the like (Mellars, 2004). But these advances have undoubtedly influenced language itself, vastly increasing the sheer number and complexity of the objects and concepts that populate our lives, which in turn may have increased the pressure to streamline and conventionalize our words and concepts, and to add complexity to sentence structure. It has also led to new forms of language—the languages of mathematics and computers. These factors, rather than the switch per se, may have altered the shape of language. In this sense, then, what linguists have generally understood as “language” may be the product of a revolution that may have been brought about through the emergence of autonomous speech, although signed languages can be easily and effectively improvised to substitute for it.

But as Tomasello (2003b) has pointed out, linguists' conceptions of language have been dominated by the languages of literate, Western populations. Across the world, languages may vary as much as the material cultures themselves do. In non-Western societies, with relatively few material artefacts, language may take a rather different shape, as exemplified by the Pirahã, but is nonetheless finely tuned to the needs and customs of the culture. It is nonetheless language, characterized by generativity and freedom of expression. Prior to the emergence of autonomous speech, a largely gestural form of language would presumably have served almost as well, but for the psychological (rather than linguistic) disadvantages of the visual modality relative to the auditory one.

So what of mirror neurons? It is clear that the discovery of mirror neurons provided strong support for the theory that language evolved from manual gesture rather than from primate calls. The mirror system in primates seems to provide a natural platform for the subsequent evolution of an intentional communicative system in which inputs are readily mapped onto outputs. In particular, one might readily speculate as to how a gestural language, perhaps eventually resembling modern signed languages, might have emerged from the basic properties of the primate mirror system. The finding that intransitive actions join transitive actions in the human mirror system paved the way to the use of the mirror system in communicative acts. The emergence of spoken language requires the further insight that speech itself can be regarded as a gestural system, but additional steps are required to account for the incorporation of vocalization into the mirror system, along with anatomical changes to provide the necessary flexibility of vocal signaling.

Of course, there must be more to language than can be understood in terms of mirror neurons themselves, or even of the extended mirror system. Even so, it is remarkable how closely the language areas of the human brain map on to the extended mirror system as identified in the primate brain (cf. Iacoboni & Wilson, 2006). This raises the question of whether the components of the mirror system have simply grown more complex to accommodate the added complexities of language, or whether brain areas unrelated to the mirror system are involved. Is language at root a matter of grasping? Thus the anterior component of the mirror system may have differentiated and elaborated to enable syntax, the posterior components to apprehend context and more subtle intention. Associative principles must also be incorporated. The mirror neurons in the monkey that fire to the sounds of actions such as tearing paper and those in humans that fire to words describing movements, where there is no somatotopic mapping, clearly imply complex learning.

Finally, language could not have evolved in a vacuum. I suggested that one capacity, possibly unique to humans, that may have driven the evolution of language was mental time travel—the ability to reconstruct past events and imagine future ones, and even to construct entirely fictional ones. Language may have evolved to allow our mental wanderings to be shared, to the benefit of the group as well as of the individual. The generativity of language and the emergence of linguistic devices to signal time and space seem precisely tuned to the requirements of communicating events, real or imaginary. This is not to say that mental time travel preceded language in evolution. My guess is that they co-evolved, to comprise the distinctive structure of the human mind.

## References

- Arbib, M. A. (2005a). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–168.
- Arbib, M. A. (2005b). Interweaving protosign and protospeech: Further developments beyond the mirror. *Interaction Studies*, 6, 145–171.
- Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language*. Cambridge: Cambridge University Press.
- Armstrong, D. F., & Wilcox, S. E. (2007). *The gestural origin of language*. Oxford: Oxford University Press.
- Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., & Mazziotta, J. (2004). Left hemisphere motor facilitation in response to manual action sounds. *European Journal of Neuroscience*, 19, 2609–2612.
- Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2006). Lateralization of the human mirror system. *Journal of Neuroscience*, 26, 2964–2970.
- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16, 1818–1823.
- Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, 40, 293–310.
- Bates, E., & Snyder, L. S. (1987). The cognitive hypothesis in language development. In E. Ina, C. Uzgiris, & E. J. McVicker Hunt (Eds.), *Infant performance and experience* (pp. 168–204). Urbana, IL: University of Illinois Press.
- Bauer, R. H. (1993). Lateralization of neural control for vocalization by the frog (*Rana pipiens*). *Psychobiology*, 21, 243–248.
- Bernardis, P., Bello, A., Pettenati, P., Stefanini, S., & Gentilucci, M. (2008). Manual actions affect vocalizations of infants. *Experimental Brain Research*, 184, 599–603.
- Bickerton, D. (1995). *Language and human behavior*. Seattle, WA: University of Washington Press.
- Binkofski, F., & Buccino, G. (2004). Motor functions of the Broca's region. *Brain and Language*, 89, 362–389.
- Bloom, P. (2004). Can a dog learn a word? *Science*, 304, 1605–1606.
- Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, 15, 19–30.
- Boë, L. J., Heim, L.-J., Honda, K., & Maeda, S. (2002). The potential neandertal vowel space was as large as that of modern humans. *Journal of Phonetics*, 30, 465–484.
- Broca, P. (1861). Remarques sur le siège de la faculté de la parole articulée, suivies d'une observation d'aphémie (perte de parole). *Bulletin de la Société d'Anatomie (Paris)*, 36, 330–357.
- Browman, C. P., & Goldstein, L. F. (1995). Dynamics and articulatory phonology. In T. van Gelder & R. F. Port (Eds.), *Mind as motion* (pp. 175–193). Cambridge, MA: MIT Press.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38.
- Burling, R. (1999). Motivation, conventionalization, and arbitrariness in the origin of language. In B. J. King (Ed.), *The origins of language: What human primates can tell us*. Santa Fe, NM: School of American Research Press.
- Burling, R. (2005). *The talking ape*. New York: Oxford University Press.
- Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience*, 15, 57–70.
- Chen, J.-Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*, 104, 427–436.
- Cheney, D. L., & Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review*, 22, 135–159.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31, 489–558.
- Christiansen, M. H., & Kirby, S. (2003). Language evolution: The hardest problem in science? In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 1–15). Oxford: Oxford University Press.

- de Condillac, E. B. (1971). *An essay on the origin of human knowledge*. T. Nugent (Tr.), Gainesville, FL: Scholars Facsimiles and Reprints (originally published 1746).
- Coop, G., Bullaughev, K., Luca, F., & Przeworski, M. (2008). The timing of selection of the human FOXP2 gene. *Molecular Biology & Evolution*, 25, 1257–1259.
- Corballis, M. C. (1991). *The lopsided ape*. New York: Oxford University Press.
- Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Corballis, M. C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, 26, 198–208.
- Corballis, M. C. (2004a). FOXP2 and the mirror system. *Trends in Cognitive Sciences*, 8, 95–96.
- Corballis, M. C. (2004b). The origins of modernity: Was autonomous speech the critical factor? *Psychological Review*, 111, 522–543.
- Corballis, M. C., & Suddendorf, T. (2007). Memory, time, and language. In C. Pasternak (Ed.), *What makes us human* (pp. 17–36). Oxford, UK: OneWorld Publications.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W.W. Norton.
- Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain*, 120, 1763–1777.
- DeGusta, D., Gilbert, W. H., & Turner, S. P. (1999). Hypoglossal canal size and hominid speech. *Proceedings of the National Academy of Sciences (USA)*, 96, 1800–1804.
- Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.
- Ehert, G. (1987). Left-hemispheric advantage in the mouse brain for recognizing ultrasonic communication calls. *Nature*, 325, 249–251.
- Emmorey, K. (2002). *Language, cognition, and brain: Insights from sign language research*. Hillsdale, NJ: Erlbaum.
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418, 869–871.
- Evans, P. D., Mekel-Bobrov, N., Yallender, E. J., Hudson, R. R., & Lahn, B. T. (2006). Evidence that the adaptive allele of the brain size gene *microcephalin* introgressed into *Homo sapiens* from an archaic *Homo* lineage. *Proceedings of the National Academy of Sciences (USA)*, 103, 18178–18183.
- Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology*, 46, 621–646.
- Everett, D. L. (2007). *Cultural constraints on grammar in Pirahã: A reply to Nevins, Pesetsky, and Rodrigues* (2007). <<http://ling.auf.net/lingBuzz/000427>>.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608–2611.
- Fauconnier, G. (1985). *Mental spaces*. Cambridge, MA: MIT Press.
- Fecteau, S., Lassonde, M., & Théoret, H. (2005). Modulation of motor cortex excitability during action observation in disconnected hemisphere. *Neuroreport*, 16, 1591–1594.
- Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, 17, 1703–1704.
- Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A., & Suomi, S. J. (2006). Neonatal imitation in rhesus macaques. *PLoS Biology*, 4, 1501–1508.
- Fogassi, L., & Ferrari, P. F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, 16, 136–141.
- Frishberg, N. (1975). Arbitrariness and iconicity in American sign language. *Language*, 51, 696–719.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13, 361–377.
- Gärdenfors, P. (2004). Cooperation and the evolution of symbolic communication. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems* (pp. 237–256). Cambridge, MA: MIT Press.
- Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165, 664–672.
- Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews*, 30, 949–960.
- Gentilucci, M., Benussi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: A kinematic study on healthy subjects. *Journal of Neurophysiology*, 86, 1685–1699.
- Gibbs, R. W. Jr., (2006). *Embodiment and cognitive science*. New York: Cambridge University Press.
- Gibson, K. R., & Jessee, S. (1999). Language evolution and expansions of multiple neurological processing areas. In B. J. King (Ed.), *The origins of language: What nonhuman primates can tell us* (pp. 189–227). Santa Fe, NM: School of American Research Press.
- Goldin-Meadow, S., & McNeill, D. (1999). The role of gesture and mimetic representation in making language the province of speech. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind* (pp. 155–172). Oxford, UK: Oxford University Press.
- Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression and speech. *Psychological Review*, 103, 34–55.
- Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215–249). Cambridge, UK: Cambridge University Press.
- Green, E., & Howes, D. H. (1978). The nature of conduction aphasia: A study of anatomic and clinical features and of underlying mechanisms. In A. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics*. San Diego: Academic Press.
- Grodzinsky, Y. (2006). The language faculty, Broca's region, and the mirror system. *Cortex*, 42, 464–468.
- Groszer, M., Keays, D. A., Deacon, R. M. J., de Bono, J. P., Prasad-Mulcare, S., Gaube, S., et al. (2008). Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Current Biology*, 18, 354–362.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39, 350–365.
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. *PLoS Biology*, 5, 2885–2897.
- Hayes, C. (1952). *The ape in our house*. London: Gollancz.
- Heine, B., & Kuteva, T. (2007). *The genesis of grammar*. Oxford: Oxford University Press.
- Hewes, G. W. (1973). Primate communication and the gestural origins of language. *Current Anthropology*, 14, 5–24.
- Hickok, G. (in press). Eight problems for mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*.
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, 15, 673–682.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the function neuroanatomy of language. *Cognition*, 92, 67–99.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4, 131–138.
- Hockett, C. (1978). In search of love's brow. *American Speech*, 53, 243–315.
- Hook-Costigan, M. A., & Rogers, L. J. (1998). Lateralized use of the mouth in production of vocalization by marmosets. *Neuropsychologia*, 36, 1265–1273.
- Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization*. Cambridge, UK: Cambridge University Press.
- Horne Tooke, J. (1857). *Epea pteroenta or the diversions of Purley*. London.
- Horvitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, K., Zilles, K., et al. (2003). Activation of Broca's area during the production of spoken and signed language: A combined cytoarchitectonic mapping and PET analysis. *Neuropsychologia*, 41, 1868–1876.
- Hurford, J. R. (2003). The language mosaic and its evolution. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (pp. 38–57). Oxford: Oxford University Press.
- Hurford, J. R. (2004). Language beyond our grasp: What mirror neurons can, and cannot, do for language evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 297–313). Cambridge, MA: MIT Press.
- Iacoboni, M., & Mazziotta, J. C. (2007). Mirror neuron system: Basic findings and clinical applications. *Annals of Neurology*, 62, 213–218.
- Iacoboni, M., & Wilson, S. M. (2006). Beyond a single area: Motor control and language within a neural architecture encompassing Broca's area. *Cortex*, 42, 503–506.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 367–371.
- Jarvis, E. D. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science*, 5, 5–14.
- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26, 235–238.
- Kaminsky, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science*, 304, 1682–1683.
- Kay, R. F., Cartmill, M., & Barlow, M. (1998). The hypoglossal canal and the origin of human vocal behavior. *Proceedings of the National Academy of Sciences (USA)*, 95, 5417–5419.
- Kendon, A. (1988). *Sign languages of aboriginal Australia*. Melbourne: Cambridge University Press.
- Kingsley, M. (1897/1965). *Travels in West Africa, Congo Française, Corisco and Cameroons*. London: F. Cass (originally published 1897).
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297, 846–848.
- Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., et al. (2007). The derived FOXP2 variant of modern humans was shared with Neanderthals. *Current Biology*, 17, 1908–1912.
- Kumashiro, M., Ishibashi, H., Uchiyama, Y., Itakura, S., Murata, A., & Iriki, A. (2003). Natural imitation induced by joint attention in Japanese macaques. *International Journal of Psychophysiology*, 50, 81–99.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Langacker, R. W. (1987). *Theoretical prerequisites. Foundations of cognitive grammar* (Vol. 1). Stanford, CA: Stanford University Press.
- Langacker, R. W. (1991). *Descriptive application. Foundations of cognitive grammar* (Vol. 2). Stanford, CA: Stanford University Press.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.

- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Liebal, K., Call, J., & Tomasello, M. (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, 64, 377–396.
- Lieberman, D. E. (1998). Sphenoid shortening and the evolution of modern cranial shape. *Nature*, 393, 158–162.
- Lieberman, D. E., McBratney, B. M., & Krovitz, G. (2002). The evolution and development of cranial form in *Homo sapiens*. *Proceedings of the National Academy of Sciences (USA)*, 99, 1134–1139.
- Lieberman, P. (1998). *Eve spoke: Human language and human evolution*. New York: W.W. Norton.
- Lieberman, P. (2007). The evolution of human speech. *Current Anthropology*, 48, 39–46.
- Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the new-born, adult human, Neanderthal man, and the chimpanzee. *American Anthropologist*, 74, 287–307.
- Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature Neuroscience*, 6, 1230–1237.
- Lin, J.-W. (2005). Time in a language without tense: The case of Chinese. *Journal of Semantics*, 23, 1–53.
- Lindenbergh, R., Fangerau, H., & Seitz, R. J. (2007). “Broca’s area” as a collective term? *Brain and Language*, 102, 22–29.
- Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13, 110–114.
- MacLarnon, A., & Hewitt, G. (2004). Increased breathing control: Another factor in the evolution of human language. *Evolutionary Anthropology*, 13, 181–197.
- MacNeillage, P. F. (2008). *The origin of speech*. Oxford: Oxford University Press.
- Mallery, G. (1880). *A collection of gesture-signs and signals of North American Indians with some comparisons*. Washington, DC: Government Printing Office.
- Masataka, N. (2001). Why early linguistic milestones are delayed in children with Williams syndrome: Late onset of hand banging as a possible rate-limiting constraint on the emergence of canonical babbling. *Developmental Science*, 4, 158–164.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Mellars, P. A. (2004). Neanderthals and the modern human colonization of Europe. *Nature*, 432, 461–465.
- Mellars, P. A., & Stringer, C. B. (Eds.). (1989). *The human revolution: Behavioral and biological perspectives on the origins of modern humans*. Edinburgh: Edinburgh University Press.
- Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American sign language*. Cambridge, MA: The MIT Press.
- Neisser, U. (2008). Memory with a grain of salt. In H. H. Wood & A. S. Byatt (Eds.), *Memory: An anthology* (pp. 80–88). London: Chatto and Windus.
- Nevins, A., Pesetsky, D., & Rodrigues, C. (2007). *Pirahã exceptionalism: A reassessment*. <<http://ling.auf.net/lingBuzz/000411>>.
- Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences (USA)*, 97, 913–918.
- Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron*, 36, 1211–1220.
- Nowak, M. A., Komorova, N. L., & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, 417, 611–617.
- Nowak, M. A., Plotkin, J. B., & Jansen, V. A. A. (2000). The evolution of syntactic communication. *Nature*, 404, 495–498.
- Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, 24, 190–198.
- Petrides, M., Cadoret, G., & Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca’s area. *Nature*, 435, 1325–1328.
- Pietrandrea, P. (2002). Iconicity and arbitrariness in Italian sign language. *Sign Language Studies*, 2, 296–321.
- Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, and use. *American Journal of Primatology*, 60, 95–111.
- Pika, S., Liebal, K., & Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): Repertoire and use. *American Journal of Primatology*, 65, 39–61.
- Pinker, S. (2007). *The stuff of thought*. London: Penguin Books.
- Pizutto, E., & Volterra, V. (2000). Iconicity and transparency in sign languages: A cross-linguistic cross-cultural view. In K. Emmorey & H. Lane (Eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima* (pp. 261–286). Mahwah, NJ: Lawrence Erlbaum.
- Pollick, A. S., & de Waal, F. B. M. (2007). Apes gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104, 8184–8189.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London B*, 358, 435–445.
- Pulvermüller, F., Huss, M., Kheri, F., de Martin, F. M. P., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103, 7865–7870.
- Ramachandran, V. S. (2000). Mirror Neurons and imitation learning as the driving force behind “the great leap forward” in human evolution. *Edge*, 69(29).
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, 21, 188–194.
- Rizzolatti, G., Camardi, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71, 491–507.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews*, 2, 661–670.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford: Oxford University Press (originally published in Italian in 2006).
- Savage-Rumbaugh, S., Shanker, S. G., & Taylor, T. J. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.
- de Saussure, F. (1916). *Cours de linguistique générale*, ed. C. Bally and A. Sechehaye, with the collaboration of A. Riedlinger, Lausanne and Paris: Payot; translated as W. Baskin (1977), *Course in general linguistics*. Glasgow: Fontana/Collins.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events. *Annals of the New York Academy of Sciences*, 1124, 39–60.
- Sekiya, K., Miyachi, S., Imaruoka, T., Egusa, H., & Tashiro, T. (2000). Body image as a visuomotor transformation device revealed in adaptation to reversed vision. *Nature*, 407, 374–377.
- Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305, 1780–1782.
- Smith, A. M. (2008). Protolanguage reconstructed. *Interaction Studies*, 9, 100–116.
- Studdert-Kennedy, M. (1998). The particulate origins of language generativity: From syllable to gesture. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language* (pp. 169–176). Cambridge, UK: Cambridge University Press.
- Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, 123, 133–167.
- Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30, 299–351.
- Sutton-Spence, R., & Boyes-Braem, P. (Eds.). (2001). *The hands are the head of the mouth: The mouth as articulator in sign language*. Hamburg: Signum-Verlag.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24, 1–86.
- Tomasello, M. (2003a). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2003b). Introduction: Some surprises for psychologists. In M. Tomasello (Ed.), *New psychology of language: Cognitive and functional approaches to language structure* (pp. 1–14). Mahwah, NJ: Lawrence Erlbaum.
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left-hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, 30, 1414–1432.
- Visalberghi, E., & Fragaszy, D. (1990). Do monkeys ape? In S. T. Parker & K. R. Gibson (Eds.), *‘Language’ and intelligence in monkeys and apes* (pp. 247–273). Cambridge: Cambridge University Press.
- Visalberghi, E., & Fragaszy, D. (2002). ‘Do monkeys ape?’ Ten years after. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 471–499). Cambridge, MA: MIT Press.
- Vleck, E. (1970). Etude comparative onto-phylogénétique de l’enfant du Pech-de-L’Azé par rapport à d’autres enfants néanderthaliens. In D. Ferembach (Ed.), *L’enfant Pech-de-L’Azé* (pp. 149–186). Paris: Masson.
- Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain*, 125, 452–464.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41, 989–994.
- Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language*, 101, 278–289.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701–702.
- Wray, A. (Ed.). (2002). *The transition to language*. Oxford: Oxford University Press.
- Wundt, W. M. (1921). *Elements of folk psychology*. New York: Macmillan.