

# Primate Vocalization, Gesture, and the Evolution of Human Language

Michael A. Arbib, Katja Liebal, and Simone Pika

CA+ Online-Only Material: Supplements A–D

The performance of language is multimodal, not confined to speech. Review of monkey and ape communication demonstrates greater flexibility in the use of hands and body than for vocalization. Nonetheless, the gestural repertoire of any group of nonhuman primates is small compared with the vocabulary of any human language and thus, presumably, of the transitional form called protolanguage. We argue that it was the coupling of gestural communication with enhanced capacities for imitation that made possible the emergence of protosign to provide essential scaffolding for protospeech in the evolution of protolanguage. Similarly, we argue against a direct evolutionary path from nonhuman primate vocalization to human speech. The analysis refines aspects of the mirror system hypothesis on the role of the primate brain's mirror system for manual action in evolution of the human language-ready brain.

In looking for the evolutionary roots of human speech, many researchers turned to the vocal signals of nonhuman primates (e.g., Seyfarth 1987; Snowdon, Brown, and Petersen 1982) as opposed to a “gestural origins” view of how language might have evolved. However, children use gestures for communication before their first spoken words, and adult speakers normally accompany all their speech with expressive manual gestures (cospeech gestures; McNeill 1992, 2005), while human signed languages are full-blown languages that do not use speech. Thus, any theory of language origins must address the fact that gestures form a crucial part of the human “language performance system.” Hewes (1973) argued that our ancestors were able to voluntarily control gestures long before speech emerged. Corballis (1991, 2002) suggested that manual gestures paved the way for the evolution of handedness linked to cerebral lateralization and—exploiting the “generativity” of manual action—for the evolution of human language. Armstrong and Wilcox (2007) support a crucial role for iconic gestures in language evolution and suggest that signed languages are the original and prototypical languages.

Our “modified gestural origins” theory charts a possible

evolutionary course from brain mechanisms for manual praxis (practical actions such as those involved in manipulating objects) to those supporting language. It does not deny the importance of vocalization but suggests that gesture and then pantomime offered a path to an open semantics that vocalization could not provide without this scaffolding. In what follows, it will be important to distinguish imitation of praxic actions from pantomime. In the early stages of our evolutionary scenario, imitation involves the attempt to repeat observed actions to achieve some goal with respect to an object. Pantomime, which we see, in evolutionary terms, as building on imitation, involves (in the early stages) the repetition of some of the movements of a praxic action, but without acting on an object, as a way of communicating something about the action, object, or event concerned.

Our theory is grounded in evidence from brain imaging (e.g., Grafton et al. 1996) that there is a human *mirror system for grasping*—i.e., a brain region activated for both grasping and observation of grasping—in or near Broca's area. Such findings raised the following question: Why might a mirror system for grasping be associated with an area commonly seen as involved in speech production? The fact that aphasia of signed and spoken languages may result from lesions to Broca's area (Emmorey 2002; Poizner, Klima, and Bellugi 1987) supports the view that one should associate Broca's area with *multimodal* language production rather than with speech alone. Such considerations led to the formulation of the mirror system hypothesis (Arbib and Rizzolatti 1997; Rizzolatti and Arbib 1998): the evolutionary basis for language parity (the more or less alignment between the meaning intended by the “speaker” and the meaning understood by the

---

**Michael A. Arbib** is Professor of Computer Science and Neuroscience at the University of Southern California (Los Angeles, CA 90089-2520, U.S.A. [arbib@usc.edu]). **Katja Liebal** has a research position at the Max Planck Institute for Evolutionary Anthropology, Leipzig, and is an honorary lecturer at the psychology department, the University of Portsmouth. **Simone Pika** is a lecturer in Evolutionary Psychology at the School of Psychological Sciences at the University of Manchester. This paper was submitted 14 XII 06 and accepted 4 VI 08.

“hearer”) is provided by the evolution of brain mechanisms that support language atop the mirror system for grasping, rooting speech in communication based on manual gesture.

In more detail, Arbib (2005a) argues that an ability for complex imitation unique to the human line made possible the evolution of brain mechanisms for pantomime and thence *protosign*, a system of conventional gestures used to formalize, disambiguate, and extend pantomime. It was further hypothesized that, once protosign has established an ability for the free creation of arbitrary gestures to support an open-ended semantics, the capacity to use conventionalized manual communicative gestures (protosign) and the capacity to use vocal communicative gestures (protospeech) evolved together in an expanding spiral (Arbib 2005b) to support *protolanguage* (Arbib 2008; Bickerton 2008), an open-ended multimodal communicative system. However, the communication systems of nonhuman primates lack *compositionality*, a crucial property of modern human languages. This is the notion that language gets its power not only from having an open-ended *lexicon* but also from having a *grammar* that allows words to be combined into phrases, with the results open to further combination, but also enables the hearer to infer the meaning of the overall utterance from the meaning of its parts and the constructions used to assemble them.

In this article, we will consider data from neuroscience only briefly. Instead, we address a glaring weakness of most writing on the mirror system hypothesis: too little attention is paid to research on the communication systems of nonhuman primates as a source of comparative data. This article is written to rectify this omission, especially with reference to the debate over whether the emergence of protosign did indeed provide essential scaffolding for the emergence of protolanguage.

The most debated topics in regard to the use of gestures, facial expressions, and vocalizations by nonhuman primates include (1) whether they are used intentionally or are simply side effects of emotional states, (2) how flexibly they are used (even gestures produced unintentionally may involve context specificity and audience effects), (3) whether they have an inherent meaning or whether the meaning is conveyed by the social context, (4) whether they are inherited or learned, and (5) whether they are used referentially.

The following sections review the existing literature on vocal communication, facial expressions, and gestural communication of nonhuman primates. We then compare communication systems in monkeys and apes and gestural communication in apes and prelinguistic or just-linguistic human children. Finally, we discuss the implications of these data for theories of language evolution.

## Vocal Communication of Nonhuman Primates

There are many studies on vocalizations of a range of monkey species (e.g., Gouzoules 1995; Kudo 1987; Seyfarth, Cheney,

and Marler 1980; Zuberbühler 2002), whereas studies of ape vocalizations focused mainly on chimpanzees (*Pan troglodytes*; e.g., Clark and Wrangham 1993; Crockford and Boesch 2003; Mitani and Gros-Louis 1998; Slocombe and Zuberbühler 2005a).

### *Production of Vocalizations*

Monkeys reared in social isolation produce basically all their species-typical call types from soon after birth. Although there is evidence of some flexibility in the way a given monkey vocalization is produced (for a recent review, see Hamerschmidt and Fischer 2008), no new vocal signals are invented by individuals (for a review, see Snowdon and Hausberger 1997). Cross-fostering of rhesus monkeys (*Macaca mulatta*) and Japanese macaques (*Macaca fuscata*) produces no significant changes in their species-specific vocalizations (Owren et al. 1992), while gibbon hybrids produce songs composed of phrases from both parental species (Geissmann 1984). Concerning apes, chimpanzee males in the wild show a positive association between the amount of time spent with another individual and call similarity in their “pant-hoots” (Mitani and Brandt 1994). However, males who chorus often with others produce more variable calls than individuals who chorus less often or call alone. Humans (*Homo sapiens*) also display a variety of involuntary vocal behaviors, but these are to be distinguished from speech (e.g., Burling 1993).

Higher degrees of flexibility are present in the “audience effect” of vocalizations (Tomasello and Zuberbühler 2002). For example, tamarins (*Saguinus labiatus*) produce food calls when discovering food, but the rates depend on whether other group mates are present (Caine, Addington, and Windfelder 1995). Vervet monkey females (*Cercopithecus aethiops*) adjust the rate of alarm calling depending on whether their own offspring are present, while males call more often when females are present (Cheney and Seyfarth 1985). Seyfarth and Cheney (2003), however, conclude that nonhuman primates may lack the human ability to represent the mental states of others and involve simply the recipients’ presence or absence.

Concerning great apes, Mitani and Nishida (1993) reported that male chimpanzees use pant-hoots more frequently in traveling contexts when their alliance partners are nearby. Wilson, Hauser, and Wrangham (2001) showed that, in response to the playback of the pant-hoot call of a single extragroup male, parties with three or more males consistently joined in a chorus of pant-hoots and approached the loudspeaker together, while parties with fewer adult males usually stayed silent and approached the loudspeaker less often. Slocombe and Zuberbühler (2007) reported that chimpanzees in the wild seemed to modify the acoustic structure of their screams during a severe attack if at least one listener in the audience matched or surpassed the aggressor in rank. This suggests that chimpanzees understand third-party relationships and adjust their vocal production in relation to the rank relationship of aggressor and listener. In addition, Hopkins,

Tagliatela, and Leavens (2007; but see also Hostetter, Cantero, and Hopkins 2001) found that chimpanzees in captivity were more likely to produce two so-called attention-getting sounds, the “raspberry” and the “extended grunt,” when a human was present in conjunction with a preferred food item than when either stimulus (human, food) was presented alone. They thus suggest that chimpanzees may produce these sounds voluntarily.

### Referential Use of Vocalizations

It has been argued that an animal vocalization qualifies as *referential* if the signal (*a*) has a distinct acoustic structure, (*b*) is produced in response to a particular external object or event, and (*c*) elicits a response in nearby listeners similar to that which the external object or event normally elicits (Zuberbühler 2000*b*). Point *b* asserts that referential communication is *triadic*, involving a sender, a receiver, and a third entity. To date, most of the evidence for referential signals in nonhuman primates stems from monkey species. For instance, vervet monkeys, Diana monkeys (*Cercopithecus diana*), and Campbell’s monkeys (*Cercopithecus campbelli*) all use distinct alarm calls for different predators, eliciting appropriate escape responses in other group members (Cheney and Seyfarth 1990; Zuberbühler 1999, 2001). Although the degree of context specificity varies across species (Evans 1997), these results suggest that referential communication is a widespread, perhaps universal, characteristic of primate communication. However, referential alarm calls are present in various mammals such as squirrels (*Spermophilus beecheyi*; Owings and Virginia 1978), marmots (Blumstein 1995*a*, 1995*b*), and even chickens (Gyger, Marler, and Pickert 1987). Such findings argue against the view that this form of the capacity to assign meaning to sound utterances constitutes a primate referential ability that could be pivotal to language. Language cannot rest on a small, fixed repertoire of such utterances.

The majority of referential calls in monkeys is linked to predation (Gouzoules and Gouzoules 2000; Gouzoules, Gouzoules, and Marler 1984; Hauser 1998). Chimpanzees in the wild produce vocalizations that are context specific and not limited to predation, crucial prerequisites for calls to function referentially (Crockford and Boesch 2003; Uhlenbroek 1996). In addition, captive chimpanzees use acoustically distinct “grunt” variants as response to different food preference classes (Slocombe and Zuberbühler 2006), and a playback experiment showed that a single chimpanzee seemed to use the information encoded in the calls to guide his search for food (Slocombe and Zuberbühler 2005*b*).

Several hypotheses may explain the difference between apes and monkeys. First, there may be a paucity of data rather than a lack of referential abilities in apes in the wild. Second, there may be evolutionary pressure for monkeys but not great apes to develop a repertoire of predator-specific alarm calls. Third, differences in social systems might account for differences in vocal behavior. Finally, great apes might have “lost”

a referential vocal system present in their ancestors because they might have become specialized for a different kind of referential skill based on the flexible use of manual gestural signals (see “Repertoire and Use”). These hypotheses need not be mutually exclusive. For example, the lack of high predator pressure in apes may have allowed them to develop a more flexible and open gestural communicative system not tied to predation.

### Acquisition of Vocalizations

Vocal production, vocal usage, and responses to vocalizations develop at different rates in primates, with vocal production being mostly innate, though the motor patterns change with maturation, and with usage conditions being affected more than the motor pattern by learning (Seyfarth and Cheney 1997). For example, the grunts of infant vervet monkeys differ from those of adults. Only later do the acoustic features of their grunts gradually come to resemble those of adults, with the grunts used appropriately (e.g., correct usage requires that an animal distinguish between dominant and subdominant individuals). The genetically determined acoustic structure of certain call types can also change as a consequence of changes in the social environment, as described for pygmy marmosets (*Cebuella pygmaea*; Snowdon and de la Torre 2002), chacma baboons (*Papio ursinus*; Fischer et al. 2004), and Campbell’s monkeys (Lemasson, Hausberger, and Zuberbühler 2005). Further evidence for the influence of the social environment on a genetically determined “basic pattern” of a given vocalization is that “coo” calls of rhesus monkeys are acoustically more similar within than between matriline (Hauser 1992). Furthermore, Japanese macaques show population-level differences in their use of food and contact calls (Green 1975; Sakura 1989), and population-specific “dialects” have been described for saddle-backed tamarins (*Saguinus fuscicollis*; Hodun, Snowdon, and Soini 1982) and chimpanzees (Mitani et al. 1992).

### Facial Expressions of Nonhuman Primates

Few studies address repertoires of facial expressions in nonhuman primates (Van Hooff 1962, 1967). Moreover, definitions differ widely. Facial expressions may be considered gestures (Maestripietri 1997; Zeller 1980) and also orofacial actions (Ferrari et al. 2003) if their production is connected to particular mouth movements such as “teeth chatter” or “lip smacks.” Some authors consider facial expressions a separate category of communicative means in addition to manual and bodily gestures (Liebal, Pika, and Tomasello 2006; Van Hooff 1962). To complicate matters further, facial expressions are often graded, and some are closely linked to the production of vocalizations, such as “horizontal pout face” in chimpanzees (linked to whimpering) or “full open grin” (linked to scream; Goodall 1986).

As for vocalizations, it is a matter of debate whether facial expressions are simply affective expressions of emotional states or whether they are intentional signals (see Caldecott 1986; Tomasello and Call 1997). A possible function of facial expressions is to serve as “metacommunicative” signals. For example, orangutans and chimpanzees use a “play face” when approaching others to make sure that a hitting gesture or wrestling is perceived as intention to play and not as an aggressive approach (Bekoff and Allen 1997; Chevalier-Skolnikoff 1994; Rijksen 1978).

A variety of facial expressions has been described for Old World monkeys, such as macaques and baboons (Hinde and Rowell 1962; Kummer 1957; Kummer and Kurt 1965), with some of them being present also in great apes (Van Hooff 1962, 1967). However, because few studies investigate the variability and frequency of facial expressions, it is unclear whether there are systematic differences in the use of facial expressions among great apes, gibbons, and monkeys. In addition, little is known about whether and how facial expressions are learned. Because rhesus macaques reared in isolation still produce their species-specific facial expressions, there seems to be a strong genetic component (Brandt, Stevens, and Mitchell 1971). The question of whether these facial expressions are produced voluntarily to influence others’ behavior remains open. In humans, the motor systems controlling affective facial expressions appear different from those controlling voluntary facial expressions, so perhaps only the affective system is operative in nonhuman primates (Gazzaniga and Smylie 1990; Rinn 1984). Tanner and Byrne (1993) observed a gorilla female who tried to hide her “play face” by covering it with her hand, consistent with the idea that the facial expression is less voluntary than the manual gesture. Too little is known about how facial expressions may relate to simultaneous gestures; this should be an object of future study.

## Gestural Communication in Apes

The use of manual and bodily gestures to communicate with other conspecifics has been reported for several species of nonhuman primates. Classic studies include those of Goodall (1986), Kummer (1968), and Van Hooff (1973; see also Hinde and Rowell 1962; Rijksen 1978), who provided detailed descriptions of different gestures (in addition to other communicative behaviors) used by monkeys and apes. More recent studies focus on the individual variability of gestural repertoires and the cognitive mechanisms underlying gestural communication (for a review, see Call and Tomasello 2007).

We next focus on ape gestures and show that (1) use of communicative gestures is common across species, (2) there is considerable variability in gesture repertoires from group to group, and (3) gestures are used flexibly in different contexts, with use depending on the behavior of the recipient. This flexibility seems attributable to learning. We will compare

studies on gestural communication of apes both in captivity and in the wild, including all great apes and siamangs (as representative of the small apes or gibbons). We consider behaviors to be gestures only if they serve to reach a recurrent social goal and are directed at a particular recipient (for criteria of gesture definition, see CA+ online supplement A). Manual and bodily gestures can be clustered into three signal categories—auditory, tactile, and visual—depending on the perceptual system used to receive them. Auditory gestures generate sound (but not with vocal cords) while tactile gestures include physical contact with the recipient and visual gestures generate a mainly visual effect with no physical contact.

### *Repertoire and Use*

A variety of gestures is reported for gibbons and great apes in both captive and wild populations. For siamangs, at least 20 different gestures, comprising both tactile and visual gestures, were observed in different captive groups (Liebal, Pika, and Tomasello 2004; see also Fox 1977; Orgeldinger 1999), with similar gestures such as “embrace” or “offer body part” also reported for white-handed gibbons (Baldwin and Teleki 1976; Ellefson 1974). For orangutans, another arboreal species, approximately 10 gestures are reported from wild populations (Mackinnon 1974; Rijksen 1978), and up to 30 different gestures are described in captive groups (Liebal, Pika, and Tomasello 2006; see also Becker 1984; Jantschke 1972; Maple 1980). For gorillas, little is known about gestural communication in wild populations (Fossey 1983; Schaller 1963; Schaller 1964), but captive gorillas utilize a variety of at least 30 different tactile, visual, and, particularly, auditory gestures (Pika, Liebal, and Tomasello 2003; Tanner 1998). Similar numbers are reported for captive chimpanzees (Tomasello et al. 1997, 1985; Van Hooff 1971), and Goodall (1986) mentions a repertoire of about a dozen gestures used in a wild population. Little is known about gestures of wild bonobos (Badrian and Badrian 1984; Ingmanson 1996; Kano 1980; Kuroda 1980, 1984). The few existing studies on individuals in captivity focus on either young individuals (performing around 20 gestures; Pika, Liebal, and Tomasello 2005) or gestures used in particular contexts, such as sex (variety of around 20 gestures; Savage-Rumbaugh, Wilkerson, and Bakeman 1977; Savage and Bakeman 1978). De Waal (1988) found that bonobos develop gestural repertoires (size around two dozen gestures) similar to those of chimpanzees but described functional differences between the two species. Play seems to be the dominant context for gesture use across captive apes (if offspring are present in a group), with the exception of orangutans, who gesture mostly in the food context (Call and Tomasello 2007; for some examples of gestures, see CA+ online supplement B; for more information and recent publications about primate gestures, see <http://www.primate-gesture-center.eu>).

### *Variability of Gestural Repertoires*

The numbers and kinds of gestures reported so far refer to the total observed in a particular population or group. However, gestural repertoires may vary depending on the individual's age and sex, as well as its group affiliation. For example, Tomasello et al. (1994, 1997) observed 30 gestures in two chimpanzee groups, but, on average, each individual used less than one-third of this repertoire. The number of gestures initially increases with age and then decreases again in adulthood (Tomasello et al. 1997). A similar pattern is found in other ape species, including siamangs (Call and Tomasello 2007). There are also group-specific gestures performed by the majority of individuals in one group but not in another (Pika, Liebal, and Tomasello 2003). "Offer arm with food pieces" in orangutans (Liebal, Pika, and Tomasello, 2006; for an example, see CA+ online supplement C), "arm shake" in gorillas (Pika, Liebal, and Tomasello 2003), and "punch" in bonobos (Pika, Liebal, and Tomasello 2005) are examples reported from captive groups, while "leaf clipping" (Nishida 1980) and "grooming hand clasp" (McGrew and Tutin 1978) are described as group-specific gestures of wild chimpanzees.

A higher degree of conformity between the gestures used by individuals is found in siamangs and gorillas living in small and stable groups, compared with species living in a more flexible social organization, such as orangutans (an individual-based fission-fusion system) and chimpanzees and bonobos (fission-fusion system), which exhibit a considerable variability of individual gestural repertoires both within and between groups (for a detailed overview of these results, see Call and Tomasello 2007). Although these results refer to captive groups, they seem consistent with the hypothesis that a species with more complex and negotiated social interactions should exhibit more variability in gesture use than a species living in small groups and/or a despotic social organization (Maestripietri 1999).

### *Flexibility*

We next consider audience effects, in regard not only to the presence/absence of a recipient but also to how gestures are adjusted depending on the attentional state or behavior of the recipient. In captivity, all apes use their visual gestures rarely, unless the recipient is visually attending (Call and Tomasello 2007). Surprisingly, they also perform at least half of their tactile gestures toward an attending audience, although this was still significantly less than for visual gestures. However, both wild and captive populations use tactile and—in case of the African great apes—auditory gestures to attract the attention of someone who is not looking at them (Nishida 1980; Tanner 2004; Tomasello et al. 1994). Orangutans adjust their begging gestures toward humans as a function of how well the humans respond (Kirk et al. 2003). If great apes have the choice of where to position themselves in relation to the orientation of a human experimenter to produce their ges-

tures, they walk in front of the human instead of manipulating his or her state of attention by using gestures behind him or her (Liebal et al. 2004).

### *Referential Use of Gestures*

There are few data on the referential use of gestures, and most of the existing literature concerns "pointing" gestures of captive chimpanzees (Leavens, Hopkins, and Thomas 2004). In a recent study, Pika and Mitani (2006) describe the widespread use of a gesture, the "directed scratch," in chimpanzees in the wild. This gesture involved one chimpanzee making a relatively loud and/or exaggerated scratching movement on a part of his body that could be seen by his grooming partner. In the majority of the cases, the indicated spot was groomed directly by the recipient. These observations suggest that this gesture is understood by receivers as referential (although self-referential) because it indicates a certain spot on the body and therefore creates a triadic communication (for an example, see CA+ online supplement D).

*Iconic* gestures relate to their referent by some actual physical resemblance such as a desired motion in space or the form of an action (Bates et al. 1979). Although iconic gestures have been reported in one bonobo and one gorilla (Tanner and Byrne [1996] reported that an adult gorilla male seemed to signal with his hand, arm, or head to a playmate the direction in which he wanted her to move or the action he wanted her to perform), these observations have not been observed in other groups of bonobos or gorillas.

### *Acquisition of Gestures*

Different mechanisms have been suggested for how nonhuman primates acquire their gestures during ontogeny, including genetic determination, ontogenetic ritualization, and social learning. Isolation experiments with rhesus macaques show that they still perform species-typical gestures and postures (Mason 1963), suggesting that the basic form of these communicative behaviors is genetically preprogrammed. Similarly, "chest beat" is reported for two gorillas that had never seen another gorilla perform these gestures (Redshaw and Locke 1976). Berdecio and Nash (1981) observed that chimpanzees from peer groups that have essentially no opportunity to observe older conspecifics develop many of the same play gestures as individuals from groups with more natural group composition. Thus, the production of at least some species-typical gestures seems to be due to genetic predisposition, triggered by commonly available individual learning conditions, as in Seyfarth and Cheney's (1997) model of vocal development.

However, there is good evidence that great apes can also invent or individually learn new gestures. Idiosyncratic gestures, used only by single individuals within a group and which therefore could not have been either genetically determined or socially learned, are reported for all great ape species

in captivity (Liebal, Pika, and Tomasello 2006; Pika, Liebal, and Tomasello 2003, 2005; Tomasello et al. 1997) and also for chimpanzees in the wild (Goodall 1986). Moreover, these gestures were used to achieve a certain social goal and most often caused a response of the recipient (Pika, Liebal, and Tomasello 2003). Thus, these individual gestures can be integrated into the group's gestural repertoire.

Tomasello and Call (1997) argue that the majority of novel gestures are learned via an individual learning process called ontogenetic ritualization. Here, a communicative signal is created by two individuals shaping each other's behavior in repeated instances of an interaction over time. For example, play hitting is an important part of the rough-and-tumble play of chimpanzees, and many individuals come to use a stylized "arm-raise" to indicate that they are about to hit the other and thus initiate play (Tomasello et al. 1997). Thus, a behavior that was not at first a communicative signal would become one over time. However, there *are* also group-specific gestures that are widely used within groups, leaving room for social learning to complement ontogenetic ritualization in the acquisition of some gestures in great apes in the wild and captivity. Unfortunately, there are no longitudinal studies investigating the ontogeny of gesture use in nonhuman primates, and existing data seem to indicate a mix of different mechanisms.

One needs to be cautious about overgeneralizing the patterns observed in captive versus wild individuals. However, some studies report that gestural repertoires are comparable between wild and captive individuals (for siamangs, see Fox 1977), although captive individuals, compared with wild ones, might use their gestures with a higher frequency (Kummer and Kurt 1965). Even if gestures might be less significant in the wild, it is striking that in captivity gestures can be found in such a variety and a high degree of variability in relatively small groups of apes, proving that gestures provide very flexible and effective communicative means that can develop even within short time spans.

## Comparing Communication Systems in Monkeys and Apes

Little is known about gestural communication in monkeys. Although there are a few studies describing gestures in hamadryas baboons and rhesus macaques (Hinde and Rowell 1962; Kummer 1957), the only systematic studies on monkey gestural communication concern macaques (*Macaca nemestrina*, *Macaca arctoides*, *Macaca mulatta* [Maestripiéri 1996a, 1996b, 1997, 1999], and *Macaca sylvanus* [Hesler and Fischer 2007]). Each species uses a variety of manual gestures and postures; the kinds of gestures individuals produce vary as a function of social context and rank (Maestripiéri 1999), and they are used flexibly across a number of different contexts (Hesler and Fischer 2007). However, these studies also show that facial expressions seem to be important communicative means in monkeys. Thus, although many facial expressions are shared

by monkeys and apes (Van Hooff 1962, 1967), it seems that facial expressions in monkeys are more prominent than in great apes (Liebal 2005). When comparing the kinds of gestures used in captive monkeys and apes, there seems to be a trend toward manual gestures in apes compared with the predominant use of postures in monkeys (Call and Tomasello 2007). However, more systematic studies are needed to address monkey gestural communication.

Facial expressions accompany both spoken and signed languages in humans but have never served to form a fully expressive language of their own. Thus, while facial expressions are important in communication (e.g., as expressions of emotional states) and in "modulating" language, their study is secondary to the primary aim of this article—namely, to assess whether nonhuman primate vocalizations or manual gestures are evolutionarily closer to the conventionalized symbol use of human language. Therefore, we set aside further consideration of facial expressions.

There is an ongoing debate about whether and to what extent nonhuman primate vocalizations are intentional, voluntarily controlled communicative means (Tomasello and Zuberbühler 2002). Although vocalizations seem to be largely innate, with a limited number of vocalizations in an individual's repertoire, there is flexibility in regard to the usage and comprehension of vocalizations, with some species even comprehending the calls of other species, which requires learning (Zuberbühler 2000a). In addition, there is some variation in certain calls as a function of population-specific dialects (Mitani, Hunley, and Murdoch 1999) or affiliation to a particular matriline (Hauser 1992).

Gestures are used intentionally and flexibly in the sense that they are directed to a specific recipient; they are adjusted to the behavior of the recipient, and one gesture can be used to achieve different goals in different contexts. Both gestures and vocalizations do occur as part of combinations or sequences (Crockford and Boesch 2005; Liebal, Call, and Tomasello 2004; Tanner 2004), with evidence that different types of combinations of vocalizations—but not gestures—convey different meanings. However, as we will discuss, such combinations are few and lack the generativity of the syntax of human languages.

For us, the crucial differences between gestures and vocalizations are that (1) gestural repertoires are open to incorporation of new gestures at both an individual and a population level; (2) there is a high degree of individual variability of gestural repertoires not only regarding age classes (as might also be the case for vocalizations) but also between different groups or populations, due to both ontogenetic ritualization and, to a lesser extent, social learning and emergence of idiosyncratic gestures; and (3) gestures are used to address one particular recipient rather than being broadcast, as is the case for most vocalizations. We thus argue that these gestures display a flexibility lacking in nonhuman primate vocalizations. This is supported by a recent study (Pollick and De Waal 2007) showing that gestures are used more flexibly than facial/

vocal signals in captive chimpanzees and bonobos. Homologous facial/vocal displays but not gestures were used similarly by both ape species.

## Comparing Gestural Communication in Apes and Humans

We stressed in the introduction to this article that gesture is an important part of language use in humans, with speech often accompanied by facial gestures that add emotional expression, as well as the hand movements known as cospeech gestures. Both of these facts seem to comport better with a view of language evolution that sees an important role for gesture than with one that seeks to trace a “voice-only” path from nonhuman primate vocalizations to language considered solely as speech. However, we postpone further development of this argument until “Implications for the Evolution of Language.” “Comparison of Ape Gestures with Gestures in Prelinguistic Human Children” offers a comparison of the use of gestures in apes with the gestural communication of prelinguistic or just-linguistic children—cospeech gestures are not discussed—and a perspective on gesture by distinguishing dyadic from triadic gestures. Then in “Teaching ‘Language’ to Apes,” we turn to attempts to teach “language” (speech and gestures) to apes and show that enculturation with humans gives apes a different gestural repertoire than that exhibited by wild individuals or even by individuals raised in captivity without human fostering. The discussion will therefore help us understand how much of complex behavior requires not only a brain with appropriate capabilities but also an environment that enables specific capabilities to develop in specific ways.

### *Comparison of Ape Gestures with Gestures in Prelinguistic Human Children*

Gestures of children can be differentiated with respect to the direction and type of gesture used (Bates 1976). The direction of gestures includes both dyadic and triadic interactions. *Dyadic* gestures are exchanged between two individuals and serve to attract the recipient’s attention toward the acting individual, whereas *triadic* gestures incorporate an external object or event into the interaction of two individuals and are used to attract the attention of the partner to this outside entity. Triadic gestures may be classified as referential gestures (see “Referential Use of Vocalization” for vocalizations and “Referential Use of Gestures” for gestures) and begin to appear in human children by the age of 12 months (Liszkowski et al. 2004).

The use of referential gestures has been linked with cognitive capacities, such as mental state attribution (Camaioni 1993; Tomasello 1995), because the recipient must infer the signaler’s meaning. Triadic gestures include both imperative and declarative gestures. *Imperative* gestures are used to get another individual to help in attaining a goal (Pika 2008a), whereas *declarative* gestures are used to draw another’s at-

tention to an outside entity to share attention (e.g., holding up an object and showing it; Pika 2008b). Plooij (1979, 1987) observed mother-infant dyads of wild chimpanzees and argued that only between the ages of 9 and 12.5 months does the chimpanzee infant start to initiate interactions with its mother by intentionally directing signals to her, for example, by using gestures such as “initiating tickling,” “grooming,” and “approach.” Only then does the chimpanzee infant act as if it understands its mother and conspecifics as social agents. This developmental stage marks the onset of the use of imperative gestures and the developmental shift from *perlocutionary* acts (where communication occurs only because the receiver is adept at interpreting the behavior of the signaler, in this case, the infant) to *illocutionary* acts in which the signaler directs his or her behavior toward a recipient.

As opposed to gestures of prelinguistic human children, the majority of gestures used in interactions between great apes can be defined as dyadic (Pika et al. 2005). Thus, a sender directs a certain gesture toward a particular recipient, with the gesture not involving an object or another outside entity. Some of the gestures apes use to attract the attention of others are “slapping the ground” in front of the recipient and “poking at” or “throwing” things at the desired partner when they want to initiate play (Tomasello, Gust, and Frost 1989). These gestures are of a triadic nature but draw, similarly to human utterances such as “Hey,” the attention to oneself and *not* to a third entity or an object (Pika et al. 2005). Moreover, chimpanzees do not use gesture combinations as a strategy to manipulate the attentional state of a recipient (Liebal, Call, and Tomasello 2004). Thus, the majority of gestures used between great apes in their natural communication are dyadic rather than triadic (Pika et al. 2005). Exceptions are the gestures “food-begging” (an animal holds out the hand, palm up to obtain food from another; for orangutans, see Bard 1992; for chimpanzees, see Tomasello et al. 1994), “food offer” (an animal offers food placed on its hand to another animal), and “present object” (an individual holds an object in front of another animal; Liebal, Pika, and Tomasello 2006). These gestures are clearly triadic. Another example is pointing, but as opposed to the previously mentioned gestures, it is deployed mainly by apes when interacting with humans (Leavens, Hopkins, and Bard 1996, 2005; Leavens, Hopkins, and Thomas 2004) or by language-trained apes (e.g., Gardner and Gardner 1969; Miles 1990; Patterson 1978a; Woodruff and Premack 1979). There are also few reports of pointing for conspecifics in captive and wild chimpanzees (de Waal 1982; Inoue-Nakamura and Matsuzawa 1997) and wild bonobos (Veà and Sabater-Pi 1998). It is important to note that apes in captivity usually point with their whole hand, not with their index finger (Leavens and Hopkins 1999). According to Leavens, Hopkins, and Thomas (2004), this behavior serves a communicative function rather than a mere reaching for food that is out of reaching distance, as argued by Povinelli and Davis (1994). (Note that pointing in humans varies between different cultures and is not restricted to index finger

pointing but can include pointing involving other body parts, such as lip pointing; Enfield 2001; Kita 2003).

### *Teaching “Language” to Apes*

Leavens, Hopkins, and Bard (2005) argued that pointing in captive apes is attributable to environmental influences on their communicative development. Another suggestion is that apes do not point for conspecifics because they do not have the motive to help or inform others or to share attention and information (Tomasello et al. 2005). The discussion makes clear that the communicative capacities of wild apes can be augmented by raising them with humans in part because humans respond in ways that apes do not. But what kinds of attempts have been made to teach apes to use human language? This bears on the issue of to what extent human language is a biological inheritance and to what extent it reflects the cumulative effects of a society’s history.

Attempts to teach apes to speak have failed repeatedly (e.g., Hayes and Hayes 1951; Kellog and Kellog 1933). Gardner and Gardner (1969) tried to overcome nonhuman primates’ difficulties in speech production by teaching American Sign Language (ASL) to a chimpanzee, Washoe (but see also Fouts and Budd 1979). Washoe did indeed learn a number of such signs, and this success led to similar projects with a gorilla, Koko (Patterson 1978*b*), and an orangutan, Chantek (Miles 1990). Other attempts to overcome the speech barrier were made by Premack (1976), who used plastic tokens to stand for spoken words in communicating with the chimpanzee Sarah. In addition, Rumbaugh (1977) created a visual language based on graphic symbols (lexigrams) depicted on a computerized keyboard for the chimpanzee Lana. Impressive results have come from the bonobo Kanzi, who spent the first 2.5 years of his life observing his mother, Matata, while she was interacting with humans around the computerized keyboard (e.g., Greenfield and Savage-Rumbaugh 1990; Savage-Rumbaugh and Brakke 1992; Savage-Rumbaugh, Shanker, and Taylor 1998). Kanzi learned many of the lexigrams that his mother had not, which implied that he had acquired them spontaneously by observing others without any specific training. A similar process is known from human children (Bruner 1983; Lock 1978), who also acquire most of their early linguistic abilities without explicit training but rather as a result of highly predictable, routine interactions with adults. Kanzi’s early vocabulary resembled that of human children, including names for individuals; labels for common objects; words for actions, locations, and properties; and even a few function words such as “no” and “yes.” His ability to understand English is comparable to that of a 2-year-old (but not older) human child (Savage-Rumbaugh et al. 1993), and he understands lexigrams as symbols in the sense that he uses them in absence of a particular referent and in a decontextualized manner. Referential abilities and increasing decontextualization are also reported for a sign language-trained orangutan, Chantek (Miles 1990).

In regard to the combinatorial aspect of symbolic communication, Fouts (1974) and Gardner and Gardner (1969) describe the spontaneous nonrandom combination of signs in ASL-using chimpanzees. However, this is not sufficient for grammar. The relationship between the symbols must be meaningful and reliable, a rule must specify the relations between categories of symbols across combinations, and the rules must be creative and productive (Terrace et al. 1979). About 10% of Kanzi’s utterances at age 5.5 years consisted of combinations of two or three lexigrams or a lexigram plus a gesture (Greenfield and Savage-Rumbaugh 1990), and he not only ordered actions but also invented his own rules. Greenfield and Savage-Rumbaugh argue that the capacity of Kanzi for some “grammatical” rules represents a “protogrammar,” indicating an evolutionary continuity with certain linguistic skills. However, compared with human children, apes acquire symbols at a much slower rate and also have a much smaller repertoire (Bonvillian and Patterson 1999; Greenfield and Savage-Rumbaugh 1990). In addition, many of their utterances represent requests and not statements or indicatives (Bonvillian and Patterson 1999). Rivas (2005) concludes that chimpanzees using ASL predominantly performed object and action signs, with no evidence for semantic or syntactic structure in combinations of signs. Some authors conclude that differences between the linguistic skills of nonhuman and human primates are quantitative rather than qualitative (Gibson 1990). However, “ape language” lacks the open-ended ability to build sentences hierarchically with a compositional semantics—that is, in such a way that the meaning of the sentence can be reconstructed from the meaning of its components by inferring how the sentence was put together. With this background, we can return to our discussion of the evolution of language.

## Implications for the Evolution of Language

In the introduction to this article, we noted that many researchers turned to the vocal signals of nonhuman primates as the direct basis for the evolution of human speech. However, we stressed that human language use is multimodal, so that any theory of language origins must include gestures as a crucial part of the human “language performance system.” We briefly recalled a number of theories laying out a gestural origin for human language, including our own “modified gestural origins” theory, the mirror system hypothesis. The aim of this article has been to provide a thorough review of the data on vocal, facial, and gestural communication in nonhuman primates as the basis for an examination of the light they shed on such theories and the standing of such theories with respect to the “direct path from vocalization” theories. As such, we aim for a focused analysis of just a few issues in



the debate on language origins rather than a comprehensive review.

*From Imitation to Pantomime: The Roots of Syntax*

Despite the ability of enculturated apes to combine signs together in an “utterance,” there is no evidence that they have the ability to employ syntax to create and re-create novel utterances that inherit their new, overall meaning from their parts via a compositional semantics. The current version of the mirror system hypothesis (Arbib 2005a) roots the essential difference in a change in skills in imitation. Where this hypothesis sees increased skill in imitation of manipulative skills as the basis for generalization to communicative skills, Donald (1991) places more emphasis on a general mimetic ability. In any case, we distinguish (i) “simple” imitation (shared with apes) that allows single actions to be acquired (more or less) in around 15 trials (Myowa-Yamakoshi and Matsuzawa 1999) and complex skills, such as strategies for feeding on nettles, to be acquired only after months of exposure to the behavior (Byrne 1999a; for related discussion, see Custance, Whiten, and Bard 1995; Russon 1996; Whiten 1998) from (ii) “complex” imitation (Arbib 2002), the capacity for recognizing novel actions as approximated by a combination of variants of known goal-directed actions. The overall action can thus be imitated immediately, though probably not skillfully at first. Wohlschläger, Gattis, and Bekkering (2003) enrich this characterization with their notion of *goal-directed imitation*. In their view, imitation is the result of perceiving an action in terms of a (possibly incomplete, possibly erroneous) hierarchical structuring of subgoals. Through these processes of successive approximation, a complex action can be reproduced with increasing accuracy by increased attention being paid to its subparts. This increased attention may result in a finer-scaled decomposition of the observed movement, resulting in execution of a more congruent behavior.

Complex actions, then—and producing a sentence in a modern human language is a complex action—involve subactions as well as actions, subgoals as well as goals. The human capacity for complex, goal-directed imitation exhibits two important properties: (a) humans can get a rough understanding of the structure of the behavior in a single exposure, with the overall program being “debugged” on subsequent exposures, rather than taking months, as in gorillas learning feeding behaviors (Byrne 1999b), and (b) humans can do it for all sorts of behaviors rather than a limited set of feeding behaviors. Wohlschläger, Gattis, and Bekkering (2003) attribute differences in imitative abilities across species to differences in working memory capacity. However, differences in imitative ability could also involve differences in the mechanism(s) of hierarchical decomposition of observed performances as well as working memory for the results (whether in planning or production).

The version of the mirror system hypothesis given by Arbib

(2005a) posits an evolutionary progression of neural systems from *simple imitation* of manual praxic actions (shared with the great apes but not monkeys) to *complex imitation* of manual praxic actions (unique to the human line) and only then turns to the evolution of communication via pantomime to protosign and beyond. However, the data reviewed in this article suggest that simple imitation in apes should be extended from manual praxic actions to include *manual communicative gestures* but not, to any great extent, vocalizations. (While some imitation-like behaviors have been observed in monkeys [Voelkl and Huber 2000, 2007], they seem to involve facilitation of use of effectors—e.g., the use of the mouth rather than the hand to remove a lid—rather than the imitation of specific [combinations of] actions.) In short, nonhuman primates have little or no capacity for vocal imitation, but apes can indeed acquire novel gestures through some combination of ontogenetic ritualization and a form of social learning akin to what we have called simple imitation (Byrne and Tanner 2006). However, whereas the demands of praxic action lead to inherently hierarchical behaviors as targets for imitation by behavior parsing, the use of gesture by apes offers little in the way of compound gestures—beyond those cases where an ape performs gestures one after the other when the recipient does not react.

Our analysis of the data on communication in nonhuman primates thus suggests that it is gesture rather than vocalization that extended the ability for simple imitation in apes into the communicative realm. This insight, missing in Arbib (2005a), yields the opening for greatly expanded gestural communication once complex imitation had evolved for practical manual skills. We posit an evolutionary progression of neural systems unique to the human line that build on those supporting complex imitation to yield (following the numbering adopted in Arbib 2005a): (S5a) pantomime of grasping and manual praxic actions; (S5b) pantomime of actions outside the pantomimic’s own behavioral repertoire (e.g., flapping the arms to mime a flying bird); (S5c) *protosign*, conventional gestures used to formalize, disambiguate, and extend pantomime (e.g., to distinguish “bird” from “flying”) that establish an ability for the free creation of arbitrary gestures to support an open-ended semantics; and, finally, (S6) *protolanguage*, a multimodal system created once protosign, after having achieved some critical mass, and *protospeech* (the capacity to use vocal communicative gestures) evolve together in an expanding spiral.

The argument for the utility of pantomime (Arbib 2002; Stokoe 2001) is that it provides an “open semantics,” allowing a large set of novel meanings to be communicated, especially as stage S5b expands on stage S5a. However, such pantomime is inefficient, both in the time taken to produce it and in the likelihood of misunderstanding. This motivates the postulation of stage S5c, as a group uses conventionalized signs to extend and exploit more efficiently the semantic richness opened by pantomime. Here, processes such as ontogenetic

ritualization convert elaborate pantomimes into a conventionalized “shorthand.”

### *Vocalization versus Gesture in the Evolution of Language*

We have argued that the successful use of gesture, via pantomime and an ability for complex imitation, was essential to the evolution of a protohuman brain that could support the open-ended expression of meaning by the compounding of gestures and that this capability provided the essential scaffolding for protospeech and the evolution of the human language-ready brain. However, we must now see how well this stands up against the argument that elaborations purely intrinsic to the core vocalization systems could have provided the basis for the evolution of brain mechanisms supporting speech. We focus our attention on a few of the strongest arguments for this view.

Cheney and Seyfarth (2005) argue that brain mechanisms for language evolved from the call systems of nonhuman primates without involvement of manual gesture because (i) even though nonhuman primate vocal repertoires contain several different call types that grade acoustically into one another, nonhuman primates produce and perceive their calls as more or less discretely different signals; (ii) different call types are given in different social contexts and listeners respond appropriately; and (iii) the grunts used by baboons (and probably many other nonhuman primates) differ according to the placement of vowel-like formants. However, our view is that these data may be relevant to the evolution of the articulatory apparatus rather than that of language.

Cheney and Seyfarth (2007), using controlled playback experiments, have also established that various calls are related to the social dominance hierarchy: baboons react more strongly to splices of calls of individuals who are not in the appropriate dominance relationship. Noting parallels between social structure and language structure, Cheney and Seyfarth (2005) hypothesize that enrichment of vocalizations for social communication provides the path to speech. However, recognizing the calls of individuals is a far cry (so to speak) from developing vocalized names for individuals or for expressing explicitly the relations between them. Nonhuman primates can explicitly express through their system of shared communicative signals only a small fraction of what their behavior demonstrates that they understand, whereas humans can express vastly more aspects of how and why they behave and can also communicate about situations and abstractions far beyond the here and now.

But perhaps monkey vocalizations share more properties with language than our review makes clear. Although most monkey calls are emitted with no particular order but rather as the context demands, there is some evidence for limited combinations (but not combinatory rules) in monkey vocal utterances. Zuberbühler (2002) demonstrated that in Diana monkeys in the wild, one specific call may act as a *semantic*

*modifier* of another. They respond to the “eagle” alarm call of Campbell’s monkeys with their own alarm calls. In situations that are less dangerous, Campbell’s monkeys combine the eagle alarm call with a so-called boom. Zuberbühler (2002) artificially combined Campbell’s monkeys’ booms with their eagle or “leopard” alarm calls and played them back to Diana monkeys. These calls no longer elicited alarm calls in Diana monkeys, indicating that the booms had affected the “meaning” of the subsequent alarm calls. When the booms preceded the alarm calls of Diana monkeys, however, they were no longer effective as semantic modifiers, indicating that the semantic modification is specific to Campbell’s monkey alarm calls.

Arnold and Zuberbühler (2006a, 2006b) found that male putty-nosed monkeys regularly combine “pyow” and “hack” calls into “pyow-hack” sequences, which usually consist of one, two, or three pyows followed by up to four hacks. They demonstrated that this combination is linked to specific external events, such as the imminent movement of the group. The authors argue that “combining existing calls into meaningful sequences increases the variety of messages that can be generated” and suggest that nonhuman primates can combine calls into higher-order sequences that have a particular meaning (Arnold and Zuberbühler 2006b, 303). However, there is no evidence of the compositionality essential to language—having a few sequences with a well-defined meaning does not qualify as syntax.

With this, let us return to our study of nonhuman primate gestures. We provided evidence that apes use more manual gestures than monkeys (e.g., Maestriperi 1999). However, apes do not seem to have a more varied repertoire of orofacial gestures. Indeed, monkeys (in particular macaques and baboons) have a variety of facial expressions, and some of them cannot be found in great apes (Van Hooff 1967). If we consider the close relationship between apes and humans, the increased use of the hands in apes compared with body postures and facial expressions in monkeys might be due to a shift from facial expressions (under less voluntary control) to more manual gestures (under voluntary control) over the course of evolution. Therefore, manual gestures may have played a role in the common ancestor of apes and humans as well. An alternative is that the use of manual gestures in apes presents a trait independent of the use of orofacial gestures and is irrelevant to human multimodal communicative behavior. However, because manual gestures still play an important role in human communication (McNeill 1992), this seems less likely. Thus, we argue that it is enrichment of the manual rather than the orofacial repertoire, exploiting the extended capability for complex imitation, that grounded the foundational semantics for an open-ended set of protosign utterances (stage S5c) and that the dramatic changes in vocal apparatus and its control (stage S6) were secondary developments in hominid evolution, driving rather than driven by expansion of the monkey orofacial repertoire.

Manual gestures in humans can have an effect on vocali-

zations made at the same time, thus creating certain natural vocal concomitants of manual efforts (Gentilucci et al. 2004a, 2004b). Chimpanzees execute sympathetic mouth movements during fine motor manipulations (Waters and Fouts 2002). These results suggest that the emergence of voice modulation and thus of an articulatory movement repertoire could have been associated with or even prompted by the preexisting manual action repertoire (for further discussion, see Roy and Arbib 2005). Such mechanisms may have played an evolutionary role in enabling the development of intentional vocal communication atop a base of gestural communication (Gentilucci and Corballis 2006).

With this, we leave aside considerations that may anchor stage S6 and turn to what may have been communicated in protosign. Recall that apes use their gestures mainly imperatively (Pika et al. 2005), whereas human children gesture for declarative purposes and to direct the attention of others to an outside object or event (Bates, Camaioni, and Volterra 1975; Liszkowski et al. 2004). We saw in “Teaching ‘Language’ to Apes” that the communicative capacities of apes in the wild can be augmented by raising them with humans in part because humans respond in ways that apes do not. The communication about outside entities may be central to the use of linguistic symbols by humans and therefore might represent a crucial difference between natural ape communication and the communication between humans and of apes with humans. We may also note the cognitive ability of humans to understand other persons as intentional agents with whom they may share a goal (Tomasello et al. 2005), an ability that is still highly debated in nonhuman primates (e.g., Heyes 1998; Pika and Zuberbühler 2007; Povinelli and Vonk 2003; Seyfarth and Cheney 2008; Tomasello, Call, and Hare 2003). Conversely, an ability to communicate a wider range of imperatives symbolically might have increased the need to communicate about more objects to which the desired actions should be directed. Imagine that there were protowords like the English “punch” and “kick.” The transition we are hinting at here would be discovering the general construction “hit with your *x*” and developing new words “*x*” for “hand” and “foot” and then many more besides. These *x*'s could then be used in turn with other predicates. This in turn would encourage triadic, referential communication that could then be expanded for purely declarative purposes. This propensity might also express the need to create a medium for the evaluation of social bonds in humans, to test and strengthen social relationships, and thus to share experiences as part of a social relationship. In our closest relatives, bonobos and chimpanzees, social grooming permeates virtually every aspect of social life. It might therefore represent their medium to evaluate and to invest into social relationships (Dunbar 1996; Pika 2008b). We are back to the importance that Cheney and Seyfarth (2005) ascribe to social structure for the emergence of language structure—but now with gesture rather than vocalization being the prime vehicle for this emergence.

## Acknowledgments

For lively discussion and/or comments on earlier drafts, we thank K. Arnold, D. Haun, H. Lyn, R. Seyfarth, and K. Zuberbühler. We also thank all of the anonymous reviewers for a range of helpful criticism and suggestions.

---

## Comments

### Michael C. Corballis

Department of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand (m.corballis@auckland.ac.nz). 12 VIII 08

The article by Arbib, Liebal, and Pika adds to the growing evidence and opinion that human language evolved from manual gestures and not from vocal calls. The authors have provided an especially thorough and pertinent review of primate vocalization, facial gestures, and manual gestures. Ironically, this review suggests more flexibility and learning in primate vocalizations than hitherto recognized, even to the point of giving at least some encouragement to those, such as Cheney and Seyfarth (2005), who still argue that language evolved from primate calls. Perhaps the clinching argument for the gestural hypothesis, though, is the at least moderate success in teaching great apes forms of gestural language, whether through signing or through pointing to visual symbols, compared with the almost total failure of teaching them to speak.

The gestural theory is further supported by evidence as to the nature and timing of anatomical changes in hominin evolution. According to most accounts, the hominins were distinguished from the other great apes by being bipedal (but see Thorpe, Holder, and Crompton 2007), which must surely have enhanced the possibilities of communicating through manual gesture more than it would have influenced the vocal repertoire. The early hominins, though, were facultative bipeds, and Donald (1991) proposed that the pantomimic phase probably originated with the genus *Homo* from around 2 million years ago, when bipedalism switched from facultative to obligate. At this time, too, brain size began its dramatic increase. In contrast, anatomical and neurophysiological adaptations for articulate vocalization probably arose much later and possibly only within the past 200,000 years, with the emergence of *Homo sapiens*. In a recent review, Lieberman (2007, 39) suggests an even more extreme scenario: “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.”

Some have argued that language itself emerged only in our own species (e.g., Bickerton 1995; Crow 2002). If we consider the complexity of human language, the more likely scenario is that it evolved gradually over the past 2 million years,

perhaps originating in a predominantly gestural form but gradually incorporating vocal elements, with vocal speech assuming dominance only in our own species. This, in turn, would have released the hands from involvement in vocalization, leading to an enhancement of manufacture and pedagogy. Speech rather than language per se might then explain how our species progressed to what has been termed “modernity” (e.g., Stringer 2002), perhaps through the release of the hands from primary involvement in communication. The rise of what has been termed “modernity” in our species might then be attributed not to language per se but to the emergence of autonomous speech (Corballis 2004).

The authors do not consider the question of what actually drove language beyond simple pantomime or protolanguage to compositional structure. At one point, though, they refer to communication “about situations and abstractions far beyond the here and now.” This may relate to the suggestion that grammatical language may have coevolved with mental time travel, also perhaps a uniquely human capacity (Suddendorf and Corballis 2007). Mental time travel includes episodic memory and the construction of possible future episodes, both of which have compositional structure. Language seems purpose built for the communication of episodes, even fictional ones, with its corresponding structure, mechanisms for coding time (e.g., tense), and use of symbols to represent episodic elements (actors, acts, patients, etc.; Corballis 2008). The emergence of compositional structure, I suggest, coevolved with mental time travel during the Pleistocene and preceded the emergence of speech as the dominant language mode.

---

### Chris Knight

Department of Social Sciences, Media, and Cultural Studies, University of East London, London E16 2RD, UK (chris.knight@uel.ac.uk). 26 VIII 08

Arbib, Liebal, and Pika provide an excellent—and long overdue—comparative survey of the incidence of gestural versus vocal communication in nonhuman primates. I like their proposal that the primate mirror neuron system underpinning gestural imitation played a key role in enabling language parity. I am also persuaded by their more general argument that the emergence of vocal speech in our ancestors in some way presupposed the scaffolding provided by gesture and then pantomime.

Unfortunately, the article ends rather abruptly, having offered little that merits description as an actual theory. The authors address a range of “what,” “when,” and “how” questions yet never ask the crucial question “Why?” Yes, apes in general do lack volitional control over their vocal signals. Yes, they do seem to have much greater control over their manual gestures. And, yes, manual gestures in humans “can have an effect on vocalizations made at the same time, thus creating certain natural vocal concomitants of manual efforts.” It

would therefore follow logically that one way an intelligent primate might enhance cognitive control over its vocal signaling would be by intentionally jumping around or otherwise manipulating its body so as to influence any sounds being emitted at the same time.

But all this strikes me as a strangely mechanistic approach to the theoretical difficulties—as if no ape or monkey ever thought to modulate its vocal signals by deploying the equipment it already has. There must surely be some more plausible reason why these animals in fact do not play around creatively or imaginatively with vocal communication. After all, young primates can be strikingly creative and imaginative in their playful antics. In the interests of masticatory efficiency, moreover, they possess jaws, lips, and tongues that are subject to fine motor control. Little effort is needed to activate the relevant mouth muscles. If greater signal flexibility would be adaptive, why not use such ready-made, highly efficient equipment to modulate sounds in the way humans do?

Instead of restricting ourselves to yet another description and classification of signaling modalities and corresponding mechanisms, we surely need some Darwinian thinking here. Among nonhuman primates, what selection pressures might have rendered it *adaptive* for vocal communication to be so strikingly insulated from cognitive control? What fitness advantages might accrue to an intelligent ape from its *inability* to play around with its vocal signals? Such questions cry out for an answer. If we do not even address them, we are unlikely to get far in elucidating the evolutionary relationships among primate vocalization, primate gesture, and speech-based human language.

The ability to engage in pantomime is, by definition, an ability to fake one’s bodily signals and displays. For patent fakes to be accepted as valid currency for purposes of communication, unusually high levels of social cooperation and corresponding trust must be assumed. This presents a theoretical conundrum because those primates intelligent enough to deploy such potentially deceptive strategies will also be clever enough to competitively exploit the trust presupposed by their habitual use (Knight 1998). This could explain why, despite their quite developed capacities for deploying and comprehending symbolic conventions when in captivity, nonhuman primates apparently find so little use for symbolic communication in the wild (Ulbaek 1998).

What would happen if a Machiavellian mutant monkey *did* discover that it could freely substitute one predator alarm call for another, regardless of the presence of any actual threat? Insofar as the fakes were exploited for purposes of tactical deception, they would lose their former status as reliable—hence meaningful—signals. To the extent that salient aspects of any signal can be intentionally faked, conspecifics will simply ignore those variable aspects in favor of any hard-to-fake acoustic features that might prove unintentionally significant. In a Darwinian social world, selection pressures will in this way drive signalers to persuade receivers of the reliability of

their signals by demonstrating precisely that they are *not* subject to cognitive control.

This will apply in particular to vocal signaling, which works at a distance, often in contexts that do not allow opportunities for immediate verification. Sound signals go around corners, work in the dark, operate over distances, and leave signalers free to continue with noncommunicative manual tasks. Such advantages make it especially important to protect the vocal-auditory modality from deceptive abuse. Lack of volitional control acts like the watermark on a banknote—it proves that the owner was not the printer. The need to guarantee reliability applies less to visual signals used in face-to-face interactions because such contexts generally offer little scope for abuse.

Facial and manual gesture work best at close quarters, in intimate contexts where immediate verification should be relatively easy. Opportunities for deception are correspondingly few. For example, when one chimpanzee informs a grooming partner at which point on its body it needs to be scratched (Pika and Mitani 2006), what could it possibly gain from a deceptive signal? It is surely no coincidence that nonhuman primates get closest to volitional referential signaling in those restricted social contexts that offer the least scope for deceptive abuse.

But this is precisely the theoretical problem: human language is not used primarily as an aid to ongoing physical activities such as grooming. Its distinctive function is “displaced reference”—communication about things not currently within sensory range. No mechanistic approach of the kind exemplified by Arbib and his colleagues can measure up to the challenge of explaining how this kind of language could possibly have evolved.

---

### David A. Leavens

Department of Psychology, School of Life Sciences,  
University of Sussex, Falmer, East Sussex BN1 9QH, UK  
(davidl@sussex.ac.uk). 27 VIII 08

Arbib, Liebal, and Pika have produced an ambitious and timely elaboration of the mirror system hypothesis of language origins. In their review, apes stand out among primates in their capacities for intentional communication with manual gestures. To the degree that the last common ancestor (LCA) of (the other) apes and humans displayed flexibility in manual signaling, over and above that seen in facial expressions and vocalizations, is the degree to which manual signaling gains allure as a prelinguistic substrate for the later complex story of language evolution. Their review concludes that great apes have moderately flexible, typically dyadic gestural communicative habits, whereas modern humans have extraordinary flexibility in triadic signaling in numerous domains. Here, I will elaborate on the significance of pointing and symbol acquisition by apes for reconstructing the triadic competencies of the LCA.

Arbib and colleagues frequently invoke the notion of flexibility in signaling. But what does “flexibility” mean in this context? In one sense, it means that an individual can use a number of different communicative tactics toward attainment of a goal and can use the same signal to different ends (Tomasello and Call 1997, 243). An example of this kind of flexibility is that chimpanzees communicate in different sensory modalities depending on whether an observer can see them (e.g., Hostetter, Cantero, and Hopkins 2001; Leavens et al. 2004). In this sense, signalers are seen as tacticians.

In a second sense, flexibility is a characteristic of dyads. In ontogenetic ritualization, pairs of social partners develop their own pair-specific patterns of nonverbal communication (Tomasello and Call 1997, 301). An example of this kind of dyadic flexibility occurs when infants ritualize their signals for being picked up: no longer straining with their whole bodies, bouncing rhythmically on their substrates, reaching with both arms fully extended, raising their arms while looking at the caregiver. Examples from adulthood include dancing and love-making—in these and other contexts, couples develop dyad-specific repertoires of nonverbal signals. Here, signalers are seen as dancers.

In a third sense, flexibility means that signals are acquired during development and shaped through common social experiences of signal consequences (Tomasello and Call 1997, 243). Whether humans typically point with their lips or fingers, for example, is a function not so much of individual tactics or dyadic accommodations but of cultural traditions (Wilkins 2003). This flexibility is manifested in patterns of within-cultural similarities as contrasted with systematic cross-cultural differences. Here, signalers are seen as being plastic in an ontogenetic sense.

According to many theorists, including Arbib and others, this third kind of ontogenetic flexibility requires sophisticated mimetic abilities. However, consideration of the distribution of pointing among apes reveals a striking pattern of group differences: wild apes almost never point manually, about half of laboratory apes point (usually with their whole hands), and virtually all language-trained apes point (usually with their index fingers; Leavens, Hopkins, and Bard 2005, 2008). Arbib (2005a, 150) noted that “chimpanzees raised in a human environment can exhibit far more ‘protolanguage’ than their wild cousins—observing animals in the wild does not define the limits of complexity of their behavior.” Apes raised by humans assume the mannerisms of their caregivers to some degree: to a lesser degree among captive apes with relatively infrequent daily interactions with caregivers and to a much greater degree among home-raised or language-trained apes.

The significance of these empirical patterns is clear when one considers that virtually everything we know about human communicative development derives from the study of humans raised in artificial urban or suburban caregiving environments that are without precedent in the prehistory of our species (e.g., Leavens, Hopkins, and Bard 2005, 2008). There is a widespread contemporary belief that, somehow, “natural”

patterns of human parenting survive, in essence, through the maelstrom of cultural evolution that characterizes a kaleidoscope of human histories. But all appeals to essence are logically antagonistic to the practice of natural history, so the mirror system hypothesis can benefit only if we acknowledge how unrepresentative any contemporary human rearing environment is of those of our prelinguistic ancestors. When we study communicative development in our own species, we see particular ontogenetic patterns only in contemporary contexts that are massively contingent on specific (pre)historical events, both biologically and culturally. The flexibility in signal repertoires developed by apes in diverse rearing environments highlights the relevance of animal models for understanding human communicative development (Bard and Leavens 2008). In particular, every one of the many thousands of linguistic signs acquired by language-trained apes entails a triadic, referential episode. These animals correctly solved the binding problem hundreds or thousands of times, and they can be characterized as habitually triadic communicators who belie classifications of apes as predominantly dyadic signalers. Human environments foster triadic communication in both humans and apes.

---

#### Dario Maestriperi

Department of Comparative Human Development,  
University of Chicago, 5730 South Woodlawn Avenue,  
Chicago, IL 60637, U.S.A. (dario@uchicago.edu).  
26 VIII 08

Many anatomical, physiological, behavioral, and cognitive traits possessed by modern humans have been produced by natural selection through a gradual process of modification of biological structures and their functional properties over a long period of time. The fact that the same traits, or similar ones, are also possessed by closely related species of apes and monkeys suggests that these traits originated long before the birth of our species, *Homo sapiens*. Our ability to use and understand language likely originated through the same natural selection process that produced other human adaptive traits. Whether language abilities are unique to humans or also shared by other primate species depends in large part on how language is defined.

Equating language with speech would make language unique to humans and reduce the usefulness of comparative primate research in understanding language evolution. Speech clearly evolved after the evolutionary split between *H. sapiens* and the other apes, so that studying the extant nonhuman primate species would not tell us much about why and how language evolved. Defining language in a way that emphasizes its relation with thought rather than communication would also reduce the usefulness of comparative primate research. Complex human thought processes were likely made possible by the substantial increase in brain size and complexity that occurred during the evolution of our hominid lineage. If lan-

guage originated as a by-product of hominid large brain size and complex cognitive skills, studying the communicative behavior of other primate species will be of limited value in understanding language evolution.

Comparative research with nonhuman primates will be most useful to understanding language evolution if language is viewed as a complex system of communication. In this case, comparative research can potentially make two types of contributions. First, it can help us identify the environmental (e.g., social or ecological) selective pressures responsible for the evolution of complex communication and possibly also the pressures responsible for the evolution of particular functional properties of language. Second, comparative research could help us identify possible precursors of human language in the communication systems of the ancestors we share with extant primate species. For example, human language may have evolved from the vocalizations of our primate ancestors, which presumably resembled those of extant nonhuman primates, or from their manual gestures, or both.

I am sympathetic to definitions of language that emphasize its functional properties and its evolutionary continuity with animal communication systems. I do believe, however, that the main contribution of comparative primate research is in the identification of the selective pressures driving the evolution of complex communication rather than the identification of language precursors in other animals' signals. The question of whether language precursors can be found among nonhuman primate vocalizations or gestures cannot be addressed with standard scientific methods and is therefore destined to remain unanswered. Moreover, neither the vocalizations nor the gestures of any extant nonhuman primates show any clear evidence of an incipient process of evolutionary modification that may have turned these signals into language. As pointed out by Arbib, Liebal, and Pika, primate vocalizations are no more complex than the vocalizations of other mammals or birds in terms of their structure, functional properties, or usage. The manual gestures of the great apes are equally unimpressive. Most, if not all, of the spontaneous gestures observed in naturalistic studies of ape behavior involve arm or hand movements displayed by juveniles or subadults in the context of social play. Although great apes exposed to humans in captivity acquire new humanlike gestures, these gestures are limited largely to requests for food. Arguing for a direct evolutionary continuity between human language and the play signals of juvenile apes or the food-begging behavior of captive apes seems unwarranted.

The mirror system hypothesis presented by Arbib and colleagues argues that the brain mechanisms for multimodal language production evolved from the mirror neuron system for grasping, a brain region activated for both grasping movements and observation of grasping. According to this hypothesis, increased skill in imitation of manipulative skills such as grasping led to an evolutionary increase in the role of imitation in the acquisition and transmission of manual gestures, which in turn resulted in greater complexity of com-

munication and, ultimately, in language. Although the hypothesis that language evolved from the mirror neuron system is plausible, the hypothesized evolutionary links between manipulative skills and manual gestures and between manual gestures and language are not directly supported by the primate behavioral data reviewed by Arbib and colleagues. The observations that social play between immature great apes involves hand gestures and that great apes can potentially use their hands to beg for food are interesting. However, the argument that these gestures may be associated with the mirror neuron system and served as evolutionary precursors of human language is far from compelling.

---

**Joanne E. Tanner**

3071 Dover Drive, Santa Cruz, CA 95065, U.S.A.  
(gorilvr@cruzio.com). 15 VIII 08

Like Arbib, Liebal, and Pika, I find that the existence of a mirror neuron system and knowledge of its workings have supported our understanding of the gestural aspect of language. However, more exploration is needed of terms central to the authors' conclusions: *imitation*, *pantomime*, and *gesture*.

Definitions of gesture in apes are not only not in agreement between investigators but also have morphed within the same research teams in their ongoing series of papers. Here, the physical nature of "gesture" goes undescribed, and the authors simply say "we consider behaviors to be gestures only if they serve to reach a recurrent social goal and are directed at a particular recipient." This statement assumes that the observer can identify an animal's social goals—which are quite possibly in constant flux and subject to changing circumstances, context, and unknown previous social history. There have been recent attempts to compare gesture repertoires of the same species (Pika 2007), but because of disparity in definitions, synthesis is far from complete. Gesture for some scholars is any kind of body movement at all and for others only "intentional" movement of the upper limbs or actions lacking mechanically effective force. Actions incorporating objects, held or thrown, are gesture in some studies and excluded by others. Some lump gesture types in nomenclature; others slice in fine detail.

A focus in my own studies has been the iconic aspect of some gesture. Borrowing from the authors' definition of pantomime, iconic gesture involves the repetition of parts of practical action without acting on objects, as communication. The authors categorize pantomime of grasping and manual praxic actions as part of an "evolutionary progression of neural systems unique to the human line." I would not want to relegate pantomime exclusively to the human line. Example: a male gorilla reaches out and taps a female on the shoulder, nods his head, and swings the arm that touched her to between his legs. The consistent result of such sequences was initiation of contact by the female (Tanner and Byrne 1999; see also Savage-Rumbaugh, Wilkerson, and Bakeman 1977

and examples from symbol-taught apes). The iconic gestures I describe in gorillas are not classic "intention movement"; they minimize force and appear to portray intentions for another animal, not the gesturer. This kind of gesture, understood instantly by others, is exactly what one would expect in a mirror system. There are also hints of "compositionality" in ape phrases such as the above. Gorillas negotiate mutual activity with gestures that function disparately as attention getters, motion depicors, location indicators, or negations (see Tanner 2004; <http://www.gorillagesures.info/>).

Facial expression, yet another form of gesture (Armstrong, Stokoe, and Wilcox 1995), needs to be studied as found within streams of gestural action. The facial expressions of gorillas in particular are an untapped area. Though they have been called less flexible than those of chimpanzees, my videotape suggests that they can be rich and varied. Further, gorillas may be aware, if not completely in control, of the communicative significance of the face (Tanner and Byrne 1993).

Early-stage imitation is distinguished from pantomime as involving the achievement of goals regarding objects. The authors also make a distinction between "simple" and "complex" (human-only) imitation that appears to be their own, though "simple imitation" seems to correspond with Byrne's (1999a, 2003) "behaviour parsing" and "complex" with Gergely, Bekkering, and Kiraly's (2002) "rational imitation" of children. Simple imitation requires repeated viewing of actions over time; complex imitation is immediate but not always wholly accurate. Thus, both categories allow practice effects, and there is overlap in categories. Tomasello et al. (1993) found enculturated chimpanzees highly successful overall in immediately reproducing both means and ends of two-step functional actions on objects with which they had been only briefly familiarized. Data on rehabilitant orangutans describe previously unfamiliar camp activities that were copied almost immediately. Though many reproductions occurred after multiple viewings of common activities, one orangutan immediately copied a rare event, axe sharpening. In an experimental setting, an orangutan copied a novel compound hierarchical action on a first trial (Russon 1996, 1999).

Complex imitation is attributed to a solely human ability to recognize a compound action as a goal-directed whole. The statement "complex imitation unique to the human line made possible the evolution of brain mechanism for pantomime" seems, in light of this discussion, (literally) out of order. Ape pantomime with a goal of movement for a social partner could have primed the brain's capacity to hold in memory longer and longer sequences of action and led to increasingly complex imitative skills.

The authors review skills of human-enculturated apes but drop them from their discussion of evolutionary progression of neural systems. Tanner, Patterson, and Byrne (2006, 87) have applied Arbib and colleagues' proposed stages of evolutionary development of language to performance of sign-taught gorilla Koko in terms of her invented (not taught) signs and found that she progresses into stages reserved. Ape

potential, like that of humans, is not limited to behavior found in native settings but influenced by the culture of their upbringing; learned or invented behavior should be recognized as part of their essential ape-ness rather than untypical.

---

## Reply

Our article surveyed data on communication systems in present-day nonhuman primates as a basis for assessing the relative merits of a “vocalization only” versus an “essential role of gesture” path from the communicative systems of our common ancestors to human language. We concluded that ape gestures show much more variation from group to group than do primate vocalizations and suggested that this ability to create new communicative signals within groups favors the latter view. In addition, we briefly discussed the pros and cons of a specific evolutionary model (the mirror system hypothesis) associated with this view. It goes far beyond what can be learned from comparative primatology but is nonetheless consistent with the constraints of data sets from a variety of disciplines.

We thus divide this reply to our commentators into two parts. The first part focuses on data on nonhuman primates, whereas the second part turns to language evolution considered more generally.

*Analysis and Interpretation of Data on Nonhuman Primates.* Tomasello and Zuberbühler (2002) proposed that vocalizations are used usually in “evolutionarily urgent” contexts such as predation or aggression and therefore must be specific in their meaning. However, as Corballis noted, we ascribe more flexibility and learning to primate vocalizations than hitherto recognized. This is in part because researchers of primate vocalizations have focused mainly on alarm calls and only now begin to investigate, for example, contact calls and food calls. Even factoring in these new data, we still see the evidence as favoring the “essential role of gesture” model—but not a “gesture only” model. We also want to stress that it is important to integrate research on different modes of communication instead of focusing on just one modality alone.

The only potentially referential gesture described so far for apes in the wild is used during grooming (Pika and Mitani, forthcoming), a context suggested as the precursor to human gossip and thus language (Dunbar 1996). Knight counters that “human language is not used primarily as an aid to ongoing physical activities such as grooming” but as displaced reference. While we agree that language today is used to direct the attentional and mental states of receivers to outside entities, its evolutionary precursors may have aided such ongoing activities. As group size increased (primate social groups ~50–55 individuals; human social network ~150 individuals), multiplying the number of relationships each individual has to monitor, servicing relationships through grooming becomes impractical (Dunbar 1996). Humans may thus have

filled the “grooming gap” by developing different communicative means—vocal and/or gestural “grooming”—which flowed later into linguistic communication (Pika 2008*b*).

Given his own contributions, Maestriperi seemed surprisingly negative about the whole field of comparative primatology of communication. Moreover, he dismissed the majority of gestures studied in apes communicating with conspecifics as juvenilia. However, we (Liebal, Pika, and Tomasello 2004, 2006) found that both young and adult siamangs and orangutans use a range of different gestures, and although all gesture types are found during play, they are also used in other contexts. Gestural repertoires decrease with age, but adults may use “new” gestures not observed in youngsters, and they may perform previously used gestures in different contexts. Further, many gestures observed in different species occur in the context of reconciliation, within the establishment or maintenance of group structure (Maestriperi 1999; Nishida et al. 1999), or during affiliative interactions and sex (Savage-Rumbaugh, Wilkerson, and Bakeman 1977; Savage-Rumbaugh and Wilkerson 1978).

Maestriperi stresses that whether language abilities are unique to humans depends in large part on how language is defined. However, we have been careful (Arbib 2005*a*) to set forth several characteristics of language that are unique to humans but are neither specific to speech nor based on the characteristics of thought.

Concerning Leavens’s discussion of pointing, the crucial question is no longer *whether* apes point (Leavens and colleagues have shown this convincingly) but *why* they point. Apes gesture mainly for imperative purposes, while humans from an early age on also gesture for declarative purposes, directing the attention of others to an outside object or event to share interest or comment on it (Bates, Camaioni, and Volterra 1975; Liszkowski 2005; Pika 2008*b*).

Tanner asks for more clarification of the term “gesture.” Moreover, she considers facial expressions as a type of gesture. While some researchers do consider facial expressions gestures (e.g., Maestriperi 1997; Zeller 1980), it remains unclear whether they are simply affective expressions of emotional states or whether they should, at least in some cases, be classified as intentional signals (e.g., Caldecott 1986; Tomasello and Call 1997). Facial expressions of emotion may be differentiated from orofacial gestures linked to ingestion (Fogassi and Ferrari 2007). However, motor linkages between arm/hand movements and speech in humans (Gentilucci, Dalla Volta, and Gianelli 2008) may have important implications for language evolution (Gentilucci and Corballis 2006).

Tanner suggests that pantomime is not exclusive to the human line and draws attention to iconic gestures, which bear some physical resemblance to their referent (Bates et al. 1979) and thus share features with pantomime (but may be highly conventionalized). However, iconic gestures have so far been described in only a single bonobo and a single gorilla (e.g., Savage-Rumbaugh, Wilkerson, and Bakeman 1977; Tanner and Byrne 1999), so the evidence is rather thin and may



hold only in the eyes of the human observer. Tanner rejects the view that complex imitation is unique to the human, arguing that ape pantomime with a goal of movement for a social partner *could* have primed the brain's capacity to hold in memory longer and longer sequences of action and led to increasingly complex imitative skills. However, given the paucity of data on nonhuman primate pantomime, the "could" here seems to us more a factor of human evolution than the development of nonhuman primates. Indeed, we distinguish the ability of apes to create and share a relatively small repertoire of gestures from the immense richness of a modern human sign language. Thus, even were it demonstrated that apes could create a few iconic gestures, these might better be seen as an evolutionary precursor to pantomime rather than pantomime in the human sense. Nonetheless, Tanner's observations make it clear that further research is needed on the comparative primatology of the diverse forms of imitation.

*General Considerations on Language Evolution.* Knight asked why vocalizations are still so limited in monkeys and explains this with mechanisms to avoid deception when communicating over larger distances and therefore being out of view. He suggests that selection pressures might have rendered it adaptive for vocal communication to be insulated from cognitive control and asks what fitness advantages might accrue to an intelligent ape from its *inability* to play around with its vocal signals. Knight sees the ability to engage in pantomime as an ability to fake one's bodily signals and displays. However, we should not conflate "honest pantomime" (or the ape gestures for which ontogenetic ritualization might be one path) with faking. Our account stresses the role of simple and then complex imitation as having a payoff in praxis that then provides opportunities for exaptation for communication. A greater emphasis on social structure would, however, usefully enrich the mirror system hypothesis by teasing out an "expanding spiral" between imitation for praxic communication and imitation for social gestures. This might in turn yield a basis for the brain structures underlying general mimesis (Donald 1998). Note, too, the advantages posited for moving from pantomime of hand gestures to pantomime of all manner of events, actions, and objects, not just actions of hominid agents as part of the move from instrumental to declarative.

Corballis asks what actually drove language beyond simple pantomime or protolanguage to compositional structure. We have argued that the ability for complex imitation enabled the evolution of brain mechanisms that provided the starting point for an open-ended spiral leading to pantomime, protosign, protolanguage, and thence compositional structure (Arbib 2008). However, we do not see language as an all-or-none system but rather suggest that any language is the result of a long process of invention, importation from other (proto)languages, and grammaticalization. Consider, then, Corballis's suggestion that the ability to travel mentally in time and space drove language from pantomime to compositional structure (Suddendorf and Corballis 2007). Certainly,

one cannot talk of the past or possible future unless one has episodic memory (which may reflect an evolutionary development from processes involving hippocampal place cells; Mizumori 2006) and the ability to weigh alternative plans of action (which, in its more sophisticated forms, may rest on integrated activity in prefrontal and parietal cortices; see, e.g., Newman 2003). However, recent studies on birds (e.g., Clayton and Dickinson 1998; Raby et al. 2007) and apes (e.g., Schwartz, Hoffmann, and Evans 2005) show that these creatures exhibit at least some aspects of episodic memory, so not all "mental time travel" is uniquely human. Moreover, it is one thing to have episodic memory or planning ability; it is quite another to communicate those memories or plans to others. We thus have a chicken and egg problem. It may be that compositionality emerged first to more flexibly describe increasingly complex episodes in the here and now (Arbib 2009) and that this made possible adding markers to descriptions of here-and-now events to indicate some form of temporal aspect. De Villiers and Pyers (2002) argue that the understanding of false beliefs supports the development of complex syntax. We have already noted the transition in human evolution from instrumental pointing to pointing for sharing interests. Any modern language is the product of bricolage, melding a large variety of different communicative and cognitive skills.

Leavens argues that the communicative development seen in humans raised in modern caregiving environments is unrepresentative of that of our prelinguistic ancestors. He also stresses the need to define the kind of flexibility involved in communicative systems. In addition to the types he mentions, flexible use can involve combining single components into larger combinations. We have suggested that the combinations seen in nonhuman primates are qualitatively different from those afforded by the grammars of human languages. This fits with our view (the mirror system hypothesis) that the brain mechanisms that support language integrate those that evolved in turn to support complex imitation, then pantomime, and then protosign and protospeech, with the richness of languages emerging much later as the result of "cultural evolution." Because most of our knowledge of triadic communication of apes stems from those raised in artificial captive environments by modern humans, it might be argued that it does not provide a useful tool to infer the communicative and cognitive skills that were available during the dawn of human language—but it does inform us about what apes are potentially capable of. Nonetheless, we are convinced that it is essential to study the impact of social and ecological factors on ape gestures in the "natural" environment as a basis for hypotheses about the role such factors might have played in the evolution of human gestural communication and language. Indeed, we emphasize the importance of combining as much as possible of the available comparative knowledge and interdisciplinary research to address the evolutionary puzzle of what makes us human.

—Michael A. Arbib, Katja Liebal, and Simone Pika

## References Cited

- Arbib, M. A. 2002. The mirror system, imitation, and the evolution of language. In *Imitation in animals and artifacts: Complex adaptive systems*, ed. K. Dautenhahn and C. L. Nehaniv, 229–80. Cambridge: MIT Press.
- . 2005a. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics (with commentaries and author's response). *Behavioral and Brain Sciences* 28:105–67.
- . 2005b. Interweaving protosign and protospeech: Further developments beyond the mirror. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems* 6:145–71.
- . 2008. Holophrasis and the protolanguage spectrum. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems* 9:151–65.
- . 2009. Invention and community in the emergence of language: A perspective from new sign languages. In *Foundations in evolutionary cognitive neuroscience: Introduction to the discipline*, ed. S. M. Platek and T. K. Shackelford, 117–52. Cambridge: Cambridge University Press.
- Arbib, M. A., and G. Rizzolatti. 1997. Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition* 29:393–424.
- Armstrong, D. F., W. C. Stokoe, and S. E. Wilcox. 1995. *Gesture and the nature of language*. Cambridge: Cambridge University Press.
- Armstrong, D. F., and S. E. Wilcox. 2007. *The gestural origin of language*. Oxford: Oxford University Press.
- Arnold, K., and K. Zuberbühler. 2006a. The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour* 72:643–53.
- . 2006b. Semantic combinations in primate calls. *Nature* 441:303.
- Badrian, A., and N. Badrian. 1984. Social organization of *Pan paniscus* in the Lomako Forest, Zaire. In *The pygmy chimpanzee: Evolutionary biology and behavior*, ed. R. L. Susman, 325–46. New York: Plenum.
- Baldwin, L. A., and G. Teleki. 1976. Patterns of gibbon behavior on Hall's Island, Bermuda: A preliminary ethogram for *Hylobates lar*. In *Gibbon and siamang*, vol. 4, ed. D. Rumbaugh, 21–105. Basel: Karger.
- Bard, K. A. 1992. Intentional behaviour and intentional communication in young free-ranging orangutans. *Child Development* 63:1186–97.
- Bard, K. A., and D. A. Leavens. 2008. Socio-emotional factors in the development of joint attention in human and ape infants. In *Learning from animals? Examining the nature of human uniqueness*, ed. L. Röska-Hardy and E. M. Neumann-Held, 89–104. London: Psychology Press. [DAL]
- Bates, E. 1976. *Language and context: The acquisition of pragmatics*. New York: Academic Press.
- Bates, E., L. Benigni, I. Bretherton, L. Camaioni, and V. Volterra. 1979. *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Bates, E., L. Camaioni, and V. Volterra. 1975. The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly* 21:205–26.
- Becker, C. 1984. *Orang-Utans und Bonobos im Spiel*. Munich: Profil.
- Bekoff, M., and C. Allen. 1997. Intentional communication and social play: How animals negotiate and agree to play. In *Animal play: Evolutionary, comparative and ecological perspectives*, ed. J. A. Byers, 97–114. Cambridge: Cambridge University Press.
- Berdecio, S., and V. J. Nash. 1981. *Chimpanzee visual communication: Facial, gestural, and postural expressive movements in young, captive chimpanzees*. Arizona State University Research Papers, vol. 26. Tempe: Arizona State University.
- Bickerton, D. 1995. *Language and human behavior*. Seattle: University of Washington Press. [MCC]
- . 2008. But how did protolanguage actually start? *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems* 9:169–76.
- Blumstein, D. T. 1995a. Golden-marmot alarm calls. 1. The production of situationally specific vocalizations. *Ethology* 100:113–25.
- . 1995b. Golden-marmot alarm calls. 2. Asymmetrical production and perception of situationally specific vocalizations. *Ethology* 101:25–32.
- Bonvillian, J. D., and F. G. Patterson. 1999. Early sign-language acquisition: Comparisons between children and gorillas. In *The mentalities of gorillas and orangutans*, ed. S. T. Parker, R. W. Mitchell, and H. L. Miles, 240–64. Cambridge: Cambridge University Press.
- Brandt, E. M., C. W. Stevens, and G. Mitchell. 1971. Visual social communication in adult male isolate-reared monkeys (*Macaca mulatta*). *Primates* 12:105–12.
- Bruner, J. S. 1983. *Child's talk*. New York: Norton.
- Burling, R. 1993. Primate calls, human language and non-verbal communication. *Current Anthropology* 34:25–53.
- Byrne, R. W. 1999a. Imitation without intentionality: Using string parsing to copy the organization of behaviour. *Animal Cognition* 2:63–72.
- . 1999b. Object manipulation and skill organization in the complex food preparation of mountain gorillas. In *The mentalities of gorillas and orangutans*, ed. S. T. Parker, R. W. Mitchell, and H. Lyn Miles, 147–59. Cambridge: Cambridge University Press.
- . 2003. Imitation as behaviour parsing. *Philosophical Transactions of the Royal Society B* 358:529–36. [JET]
- Byrne, R. W., and J. E. Tanner. 2006. Gestural imitation by a gorilla: Evidence and nature of capacity. *International Journal of Psychology and Psychological Therapy* 6:215–32.
- Caine, N. G., R. L. Addington, and T. L. Windfelder. 1995. Factors affecting the rates of food calls given by red-bellied tamarins. *Animal Behaviour* 50:53–60.

- Caldecott, J. O. 1986. *An ecological and behavioural study of the pig-tailed macaque*. Basel: Karger.
- Call, J., and M. Tomasello, eds. 2007. *The gestural communication of apes and monkeys*. Mahwah, N.J.: Lawrence Erlbaum.
- Camaioni, L. 1993. The development of intentional communication: A re-analysis. In *New perspectives in early communicative development*, ed. J. Nadel and L. Camaioni, 82–96. London: Routledge.
- Cheney, D. L., and R. M. Seyfarth. 1985. Vervet monkey alarm calls: Manipulation through shared information? *Behavior* 94:150–66.
- . 1990. *How monkeys see the world*. Chicago: University of Chicago Press.
- . 2005. Constraints and preadaptations in the earliest stages of language evolution. *Linguistic Review* 22:135–59.
- . 2007. *Baboon metaphysics: The evolution of a social mind*. Chicago: University of Chicago Press.
- Chevalier-Skolnikoff, S. 1994. The primate play face: A possible key to the determinants and evolution of play. *Rice University Studies* 60:9–29.
- Clark, A. P., and R. W. Wrangham. 1993. Acoustic analysis of wild chimpanzee pant hoots: Do Kibale forest chimpanzees have an acoustically distinct food arrival pant hoot? *American Journal of Primatology* 31:99–109.
- Clayton, N. S., and A. Dickinson. 1998. Episodic-like memory during cache recovery by scrub jays. *Nature* 395:272–78.
- Corballis, M. C. 1991. *The lopsided ape: Evolution of the generative mind*. New York: Oxford University Press.
- . 2002. *From hand to mouth, the origins of language*. Princeton: Princeton University Press.
- . 2004. The origins of modernity: Was autonomous speech the critical factor? *Psychological Review* 111:543–52. [MCC]
- . 2008. Mental time travel and the shaping of language. *Experimental Brain Research* (forthcoming). [MCC]
- Crockford, C., and C. Boesch. 2003. Context-specific calls in wild chimpanzees, *Pan troglodytes verus*: Analysis of barks. *Animal Behaviour* 66:115–25.
- . 2005. Call combinations in wild chimpanzees. *Behaviour* 142:397–421.
- Crow, T. J. 2002. Sexual selection, timing, and an X-Y homologous gene: Did *Homo sapiens* speciate on the Y chromosome? In *The speciation of modern Homo sapiens*, ed. T. J. Crow, 197–216. Oxford: Oxford University Press. [MCC]
- Custance, D. M., A. Whiten, and K. A. Bard. 1995. Can young chimpanzees (*Pan troglodytes*) imitate arbitrary actions? Hayes & Hayes (1952) revisited. *Behaviour* 132:11–12.
- de Villiers, J. G., and J. E. Pyers. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17:1037–60.
- de Waal, F. B. M. 1982. *Chimpanzee politics*. Baltimore: Johns Hopkins University Press.
- . 1988. The communicative repertoire of captive bonobos (*Pan paniscus*) compared to that of chimpanzees. *Behaviour* 106:183–251.
- Donald, M. 1991. *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge: Harvard University Press.
- . 1998. Mimesis and the executive suite: Missing links in language evolution. In *Approaches to the evolution of language: Social and cognitive bases*, ed. J. R. Hurford, M. Studdert-Kennedy, and C. Knight, 44–67. Cambridge: Cambridge University Press.
- Dunbar, R. 1996. *Grooming, gossip and the evolution of language*. London: Faber and Faber.
- Ellefsen, J. O. 1974. A natural history of white-handed gibbons in the Malayan peninsula. In *Gibbon and siamang*, ed. D. M. Rumbaugh, 2–143. Basel: Karger.
- Emmorey, K. 2002. *Language, cognition, and the brain: Insights from sign language research*. Mahwah, N.J.: Lawrence Erlbaum.
- Enfield, N. J. 2001. “Lip-pointing”: A discussion of form and function with reference to data from Laos. *Gesture* 1: 185–211.
- Evans, C. S. 1997. Referential signals. *Perspectives in Ethology* 12:99–143.
- Ferrari, P. F., V. Gallese, G. Rizzolatti, and L. Fogassi. 2003. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience* 17: 1703–14.
- Fischer, J., D. M. Kitchen, R. M. Seyfarth, and D. L. Cheney. 2004. Baboon loud calls advertise male quality: Acoustic features. *Behavioral Ecology and Sociobiology* 56:140–48.
- Fogassi, L., and P.-F. Ferrari. 2007. Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science* 16:136–41.
- Fossey, D. 1983. *Gorillas in the mist*. London: Hodder and Stoughton.
- Fouts, R. S. 1974. Capacities for language in great apes. In *Socioecology and psychology of primates*, ed. R. H. Tuttle, 371–90. The Hague: Mouton.
- Fouts, R. S., and R. L. Budd. 1979. Artificial and human language acquisition in the chimpanzee. In *The great apes*, ed. D. Hamburg and E. R. McCown, 375–92. Menlo Park, Calif.: Benjamin/Cummings.
- Fox, G. J. 1977. *Social dynamics in siamang*. Milwaukee: University of Wisconsin Press.
- Gardner, R. A., and B. Gardner. 1969. Teaching sign language to a chimpanzee. *Science* 165:664–72.
- Gazzaniga, M. S., and C. S. Smylie. 1990. Hemispheric mechanisms controlling voluntary and spontaneous facial expressions. *Journal of Cognitive Neuroscience* 2:239–45.
- Geissmann, T. 1984. Inheritance of song parameters in the gibbon song analysed in two hybrid gibbons (*Hylobates pileatus* × *Hylobates lar*). *Folia primatologica* 42:216–35.
- Gentilucci, M., and M. C. Corballis. 2006. From manual ges-

- ture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews* 30:949–60.
- Gentilucci, M., R. Dalla Volta, and C. Gianelli. 2008. When the hands speak. *Journal of Physiology* (Paris) 102:21–30.
- Gentilucci, M., P. Santunione, A. C. Roy, and S. Stefanini. 2004a. Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *European Journal of Neuroscience* 19:190–202.
- Gentilucci, M., S. Stefanini, A. C. Roy, and P. Santunione. 2004b. Action observation and speech production: Study on children and adults. *Neuropsychologia* 42:1554–67.
- Gergely, G., H. Bekkering, and I. Kiraly. 2002. Rational imitation in preverbal infants. *Nature* 415:755. [JET]
- Gibson, K. R. 1990. New perspectives on instincts and intelligence: Brain size and the emergence of hierarchical mental constructional skills. In *“Language” and intelligence in monkeys and apes*, ed. S. T. Parker and K. R. Gibson, 97–128. Cambridge: Cambridge University Press.
- Goodall, J. 1986. *The chimpanzees of Gombe: Patterns of behaviour*. Cambridge: Belknap.
- Gouzoules, H. S. 1995. Representational signaling in non-human primate vocal communication. In *Current topics in primate vocal communication*, ed. E. Zimmermann, J. D. Newman, and U. Jürgens, 235–52. New York: Plenum.
- Gouzoules, H. S., and S. Gouzoules. 2000. Agonsitic screams differ among four species of macaques: The significance of motivation-structural rules. *Animal Behaviour* 59:501–12.
- Gouzoules, S., H. Gouzoules, and P. Marler. 1984. Rhesus monkey (*Macaca mulatta*) screams: Representational signalling in the recruitment of agonistic aid. *Animal Behaviour* 32:182–93.
- Grafton, S. T., M. A. Arbib, L. Fadiga, and G. Rizzolatti. 1996. Localization of grasp representations in humans by PET. 2. Observation compared with imagination. *Experimental Brain Research* 112:103–11.
- Green, S. 1975. Variation of vocal pattern with social situation in the Japanese monkey: A field study. *Primate Behaviour* 4:1–102.
- Greenfield, P. M., and E. S. Savage-Rumbaugh. 1990. Grammatical combination in *Pan paniscus*: Processes of learning and invention in the evolution and development of language. In *“Language” and intelligence in monkeys and apes*, ed. S. T. Parker and K. R. Gibson, 540–78. Cambridge: Cambridge University Press.
- Gyger, M., P. Marler, and R. Pickert. 1987. Semantics of an avian alarm call system: The male domestic fowl, *G. domesticus*. *Behaviour* 102:15–40.
- Hammerschmidt, K., and J. Fischer. 2008. Constraints in primate vocal production. In *The evolution of communicative creativity: From fixed signals to contextual flexibility*, ed. U. Griebel and K. Oller. Cambridge: MIT Press (forthcoming).
- Hauser, M. 1992. Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production? *Journal of the Acoustical Society of America* 91:2175–79.
- Hauser, M. D. 1998. Functional referents and acoustic similarity: Field playback experiments with rhesus monkeys. *Animal Behaviour* 55:1309–10.
- Hayes, K. J., and C. Hayes. 1951. The intellectual development of a home-raised chimpanzee. *Proceedings of the American Philosophical Society* 95:105–9.
- Hesler, N., and J. Fischer. 2007. Gestural communication in Barbary macaques (*Macaca sylvanus*): An overview. In *The gestural communication of apes and monkeys*, ed. J. Call and M. Tomasello, 159–96. Mahwah, N.J.: Lawrence Erlbaum.
- Hewes, G. W. 1973. Primate communication and the gestural origin of language. *Current Anthropology* 12:5–24.
- Heyes, C. M. 1998. Theory of mind in nonhuman primates. *Behavioral and Brain Science* 21:101–48.
- Hinde, R. A., and T. E. Rowell. 1962. Communication by postures and facial expressions in the rhesus monkey, *Macaca mulatta*. *Proceedings of the Zoological Society of London* 138:1–21.
- Hodun, A., C. T. Snowdon, and P. Soini. 1982. Subspecific variation in the long calls of the tamarin, *Saguinus fuscicollis*. *Zeitschrift für Tierpsychologie* 57:97–110.
- Hopkins, W. D., J. P. Tagliabue, and D. Leavens. 2007. Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour* 73:281–86.
- Hostetter, Autumn B., Monica Cantero, and William D. Hopkins. 2001. Differential use of vocal and gestural communication by chimpanzees (*Pan troglodytes*) in response to the attentional status of a human (*Homo sapiens*). *Journal of Comparative Psychology* 115:337–43.
- Ingmanson, E. J. 1996. Tool-using behavior in wild *Pan paniscus*: Social and ecological considerations. In *Reaching into thought: The minds of the great apes*, ed. A. E. Russon and K. A. Bard, 190–210. Cambridge: Cambridge University Press.
- Inoue-Nakamura, N., and T. Matsuzawa. 1997. Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 111:159–73.
- Jantschke, F. 1972. *Orang utans in Zoologischen Gärten*. Munich: Piper.
- Kano, T. 1980. Social behaviour of wild pygmy chimpanzees (*Pan paniscus*) of Wamba: A preliminary report. *Journal of Human Evolution* 9:243–60.
- Kellog, W. N., and L. A. Kellog. 1933. *The ape and the child*. New York: McGraw-Hill.
- Kirk, E. C., M. Cartmill, R. F. Kay, and P. Lemelin. 2003. Comment on “Grasping primate origins.” *Science* 300:741; author reply 741.
- Kita, S. 2003. *Pointing: Where language, culture, and cognition meet*. Mahwah, N.J.: Lawrence Erlbaum.
- Knight, C. 1998. Ritual/speech coevolution: A solution to the problem of deception. In *Approaches to the evolution of language: Social and cognitive bases*, ed. J. R. Hurford, M. Studdert-Kennedy, and C. Knight, 68–91. Cambridge: Cambridge University Press. [CK]
- Kudo, H. 1987. The study of vocal communication of wild

- mandrills in Cameroon in relation to their social structure. *Primates* 28:289–308.
- Kummer, H. 1957. Soziales Verhalten einer Mantelpavian-Gruppe. *Schweizerische Zeitschrift für Psychologie und ihre Anwendungen* 33:1–91.
- . 1968. *Social organization of hamadryas baboons*. Chicago: University of Chicago Press.
- Kummer, H., and F. Kurt. 1965. A comparison of social behaviour in captive and wild hamadryas baboons. In *The baboon in medical research*, ed. H. Vagtborg, 1–16. Austin: University of Texas Press.
- Kuroda, S. J. 1980. Social behavior of the pygmy chimpanzees. *Primates* 21:181–97.
- . 1984. Rocking gesture as communicative behavior in the wild pygmy chimpanzee in Wamba, Central Zaire. *Journal of Ethology* 2:127–37.
- Leavens, D. A., and W. D. Hopkins. 1999. The whole hand-point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology* 113: 417–25.
- Leavens, D. A., W. D. Hopkins, and K. A. Bard. 1996. Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 110:346–53.
- . 2005. Understanding the point of chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in Psychological Science* 14:185–89.
- . 2008. The heterochronic origins of explicit reference. In *The shared mind: Perspectives on intersubjectivity*, ed. Jordan Zlatev, Timothy P. Racine, Chris Sinha, and Esa Itkonen, 187–214. Amsterdam: John Benjamins. [DAL]
- Leavens, D. A., W. D. Hopkins, and R. K. Thomas. 2004. Referential communication by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 118:48–57.
- Leavens, David A., Autumn B. Hostetter, Michael J. Wesley, and William D. Hopkins. 2004. Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Animal Behaviour* 67:467–76. [DAL]
- Lemasson, A., M. Hausberger, and K. Zuberbühler. 2005. Socially meaningful vocal plasticity in adult Campbell's monkeys (*Cercopithecus campbelli*). *Journal of Comparative Psychology* 119:220–29.
- Liebal, K. 2005. Social communication in great apes. Ph.D. diss., Universität Leipzig.
- Liebal, K., J. Call, and M. Tomasello. 2004. Chimpanzee gesture sequences. *Primates* 64:377–96.
- Liebal, K., S. Pika, J. Call, and M. Tomasello. 2004. Great ape communicators move in front of recipients before producing visual gestures. *Interaction Studies* 5:199–219.
- Liebal, K., S. Pika, and M. Tomasello. 2004. Social communication in siamangs (*Symphalangus syndactylus*): Use of gestures and facial expression. *Primates* 45:41–57.
- . 2006. Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6:1–38.
- Lieberman, P. 2007. The evolution of human speech. *Current Anthropology* 48:39–66. [MCC]
- Liszkowski, U. 2005. Human twelve-month-olds point cooperatively to share interest with and provide information for a communicative partner. *Gesture* 5:135–54.
- Liszkowski, U., M. Carpenter, A. Henning, T. Striano, and M. Tomasello. 2004. Twelve-month-olds point to share attention and interest. *Developmental Science* 7:297–307.
- Lock, A. 1978. *Action, gesture and symbol: The emergence of language*. New York: Academic Press.
- Mackinnon, J. R. 1974. Behaviour and ecology of orangutans. *Animal Behaviour* 22:3–74.
- Maestripieri, D. 1996a. Gestural communication and its cognitive implications in pigtail macaques (*Macaca nemestrina*). *Behaviour* 133:997–1022.
- . 1996b. Social communication among captive stump-tailed macaques (*Macaca arctoides*). *International Journal of Primatology* 17:785–802.
- . 1997. Gestural communication in macaques. *Evolution of Communication* 1:193–222.
- . 1999. Primate social organization, gestural repertoire size, and communication dynamics. In *The origins of language: What nonhuman primates can tell*, ed. B. J. King, 55–77. Santa Fe: School of American Research Press.
- Maple, T. L. 1980. *Orangutan behavior*. New York: Van Nostrand Reinhold Primate Behaviour and Development Series.
- Mason, W. A. 1963. Social development of rhesus monkeys with restricted social experience. *Perceptual and Motor Skills* 16:263–70.
- McGrew, W. C., and C. E. G. Tutin. 1978. Evidence for a social custom in wild chimpanzees? *Man* 13:234–51.
- McNeill, D. 1992. *Hand and mind*. Chicago: University of Chicago Press.
- . 2005. *Gesture and thought*. Chicago: University of Chicago Press.
- Miles, H. L. 1990. The cognitive foundations for reference in a signing orangutan. In *“Language” and intelligence in monkeys and apes*, ed. S. T. Parker and K. R. Gibson, 511–39. Cambridge: Cambridge University Press.
- Mitani, J. C., and K. L. Brandt. 1994. Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology* 96:233–52.
- Mitani, J. C., and J. Gros-Louis. 1998. Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* 135:1041–64.
- Mitani, J. C., T. Hasegawa, J. Gros-Louis, P. Marler, and R. W. Byrne. 1992. Dialects in wild chimpanzees? *American Journal of Primatology* 27:233–43.
- Mitani, J. C., K. L. Hunley, and M. E. Murdoch. 1999. Geographic variation in the calls of wild chimpanzees: A reassessment. *American Journal of Primatology* 47:133–51.
- Mitani, J. C., and T. Nishida. 1993. Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour* 45:735–46.
- Mizumori, S. J. 2006. Hippocampal place fields: A neural code for episodic memory? *Hippocampus* 16:685–90.

- Myowa-Yamakoshi, M., and T. Matsuzawa. 1999. Factors influencing imitation of manipulatory actions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 113:128–36.
- Newman, J. D. 2003. Vocal communication and the triune brain. *Physiology and Behavior* 79:495–502.
- Nishida, T. 1980. The leaf-clipping display: A newly discovered expressive gesture in wild chimpanzees. *Journal of Human Evolution* 9:117–28.
- Nishida, T., T. Kano, J. Goodall, W. C. McGrew, and M. Nakamura. 1999. Ethogram and ethnography of Mahale chimpanzees. *Anthropological Science* 107:141–88.
- Orgeldinger, M. 1999. *Paarbeziehungen beim Siamang-Gibbon (Hylobates syndactylus) im Zoo: Untersuchungen über den Einfluß von Jungtieren auf die Paarbindung*. Münster: Schöningh.
- Owings, D. H., and R. A. Virginia. 1978. Alarm calls of California ground squirrels (*Spermophilus beecheyi*). *Zeitschrift für Tierpsychologie* 46:58–70.
- Owren, M. J., J. A. Dieter, R. M. Seyfarth, and D. L. Cheney. 1992. Evidence of limited modification in the vocalizations of cross-fostered rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques. In *Topics in primatology: Human origins*, ed. T. Nishida, W. C. McGrew, P. Marler, M. Pickford, and F. B. M. de Waal, 257–70. Tokyo: University of Tokyo Press.
- Patterson, F. 1978a. Conversations with a gorilla. *National Geographic* 134:438–65.
- . 1978b. Linguistic capabilities of a lowland gorilla. In *Sign language and language acquisition in man and ape*, ed. F. C. C. Peng, 161–201. Boulder: Westview Press.
- Pika, S. 2007. Gestures in subadult gorillas. In *The gestural communication of monkeys and apes*, ed. Josep Call and Michael Tomasello, 99–130. Mahwah, N.J.: Lawrence Erlbaum. [JET]
- . 2008a. Gestures of apes and pre-linguistic human children: Similar or different? *First Language* 28:116–40.
- . 2008b. What is the nature of the gestural communication of great apes? In *The shared mind*, ed. J. Zlatev, T. Racine, C. Sinha, and E. Itkonen, 165–86. Amsterdam: John Benjamins Publishing Company.
- Pika, S., K. Liebal, J. Call, and M. Tomasello. 2005. The gestural communication of apes. *Gesture* 5:41–56.
- Pika, S., K. Liebal, and M. Tomasello. 2003. Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, learning and use. *American Journal of Primatology* 60:95–111.
- . 2005. Gestural communication in subadult bonobos (*Pan paniscus*): Repertoire and use. *American Journal of Primatology* 65:39–61.
- Pika, S., and J. C. Mitani. 2006. Referential gesturing in wild chimpanzees (*Pan troglodytes*). *Current Biology* 16:191–92.
- . Forthcoming. The directed scratch: Evidence for a referential gesture in chimpanzees? In *The prehistory of language*, ed. R. Botha and C. Knight. Oxford: Oxford University Press.
- Pika, S., and K. Zuberbühler. 2007. Social games between bonobos and humans: Evidence for shared intentionality? *American Journal of Primatology* 70:207–10.
- Plooij, F. X. 1979. How wild chimpanzee babies trigger the onset of mother-infant play. In *Before speech*, ed. M. Bullock, 223–43. Cambridge: Cambridge University Press.
- . 1987. Infant-ape behavioural development, the control of perception and, types of learning and symbolism. In *Symbolism and knowledge*, vol. 8, ed. A. Tryphon and J. Montangero, 29–58. Geneva: Jean Piaget Archives Foundation.
- Poizner, H., E. Klima, and U. Bellugi. 1987. *What the hands reveal about the brain*. Cambridge: MIT Press.
- Pollick, A. S., and F. B. M. De Waal. 2007. Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the USA* 104:8184–89.
- Povinelli, D. J., and D. R. Davis. 1994. Differences between chimpanzees (*Pan troglodytes*) and humans (*Homo sapiens*) in the resting state of the index finger: Implications for pointing. *Journal of Comparative Psychology* 108:134–39.
- Povinelli, D. J., and J. Vonk. 2003. Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences* 4:157–60.
- Premack, D. 1976. *Intelligence in ape and man*. Hillsdale, N.Y.: Lawrence Erlbaum.
- Raby, C. R., D. M. Alexis, A. Dickinson, and N. S. Clayton. 2007. Planning for the future by western scrub-jays. *Nature* 445:919–21.
- Redshaw, M., and K. Locke. 1976. The development of play and social behaviour in two lowland gorilla infants. *Journal of the Jersey Wildlife Preservation Trust, Thirteenth Annual Report*, 71–86.
- Rijksen, H. D. 1978. *A field study on Sumatran orangutans*. Wageningen: Mededelingen Landbouwhogeschool.
- Rinn, W. E. 1984. The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin* 95: 52–77.
- Rivas, E. 2005. Recent use of signs by chimpanzees (*Pan troglodytes*) in interactions with humans. *Journal of Comparative Psychology* 119:404–17.
- Rizzolatti, G., and M. A. Arbib. 1998. Language within our grasp. *Trends in Neuroscience* 21:188–94.
- Roy, A. C., and M. A. Arbib. 2005. The syntactic motor system. *Gesture* 5:7–37.
- Rumbaugh, D. M. 1977. *Language learning by a chimpanzee: The Lana project*. New York: Academic Press.
- Russon, A. E. 1996. Imitation in everyday use: Matching and rehearsal in the spontaneous imitation of rehabilitant orangutans (*Pongo pygmaeus*). In *Reaching into thought: The minds of the great apes*, ed. A. E. Russon, K. A. Bard, and S. T. Parker, 152–76. Cambridge: Cambridge University Press.
- . 1999. Orangutans' imitation of tool use: A cognitive interpretation. In *The mentalities of gorillas and orangutans*, ed. Sue Taylor Parker, Robert W. Mitchell, and H. Lyn

- Miles, 117–46. Cambridge: Cambridge University Press. [JET]
- Sakura, O. 1989. Variability in contact calls between troops of Japanese macaques: A possible case of neutral evolution of animal culture. *Animal Behaviour* 38:900–902.
- Savage, S., and R. Bakeman. 1978. Sexual morphology and behavior in *Pan paniscus*. In *Recent advances in primatology*, ed. D. J. Chivers and J. Herbert, 613–16. New York: Academic Press.
- Savage-Rumbaugh, E., J. Murphy, R. A. Sevcic, K. E. Brakke, S. L. Williams, and D. M. Rumbaugh. 1993. Language comprehension in ape and child. *Monographs of the Society for Research in Child Development* 58:1–256.
- Savage-Rumbaugh, E. S., and K. E. Brakke. 1992. Linguistic development: Contrasts between co-reared *Pan troglodytes* and *Pan paniscus*. In *Topics in primatology: Human origins*, vol. 1, ed. T. Nishida, W. C. McGrew, P. Marler, M. Pickford, and F. B. M. de Waal, 51–66. Tokyo: University of Tokyo Press.
- Savage-Rumbaugh, E. S., S. G. Shanker, and T. J. Taylor. 1998. *Apes, language, and the human mind*. New York: Oxford University Press.
- Savage-Rumbaugh, E. S., B. J. Wilkerson, and R. Bakeman. 1977. Spontaneous gestural communication among conspecifics in the pygmy chimpanzee (*Pan paniscus*). In *Progress in ape research*, ed. G. H. Bourne, 97–116. New York: Academic Press.
- Savage-Rumbaugh, S., and B. Wilkerson. 1978. Socio-sexual behavior in *Pan paniscus* and *Pan troglodytes*: A comparative study. *Journal for Human Evolution* 7:327–44.
- Schaller, G. B. 1963. *The mountain gorilla, ecology and behavior*. Chicago: University of Chicago Press.
- . 1964. The behavior of the mountain gorilla. In *Primate behavior field studies of monkeys and apes*, ed. I. DeVore, 324–67. New York: Holt, Rinehart, and Winston.
- Schwartz, B. L., M. L. Hoffmann, and S. Evans. 2005. Episodic-like memory in a gorilla: A review and new findings. *Learning and Motivation* 36:226–44.
- Seyfarth, R. M. 1987. Vocal communication and its relation to language. In *Primate societies*, ed. B. Smuts, D. L. Cheney, R. Seyfarth, R. Wrangham, and T. Struhsaker, 440–51. Chicago: University of Chicago Press.
- Seyfarth, R. M., and D. L. Cheney. 1997. Some general features of vocal development in nonhuman primates. In *Social influences on vocal development*, ed. C. Snowdon and M. Hausberger, 249–73. Cambridge: Cambridge University Press.
- . 2003. The structure of social knowledge in monkeys. In *Animal social complexity: Intelligence, culture and individualized societies*, ed. F. B. M. de Waal and P. L. Tyack, 207–29. Cambridge: Harvard University Press.
- . 2008. Primate social knowledge and the origins of language. *Mind & Society* 7:129–42.
- Seyfarth, R. M., D. L. Cheney, and P. Marler. 1980. Monkey responses to three different alarm calls: Evidence for predator classification and semantic communication. *Science* 210:801–3.
- Slocombe, K. E., and K. Zuberbühler. 2005a. Agonistic screams in wild chimpanzees vary as a function of social role. *Journal of Comparative Psychology* 119:67–77.
- . 2005b. Functionally referential communication in a chimpanzee. *Current Biology* 15:1779–784.
- . 2006. Food-associated calls in chimpanzees: Responses to food types or food preferences? *Animal Behaviour* 72:989–99.
- . 2007. Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences of the USA* 104:17228–33.
- Snowdon, C. T., C. H. Brown, and M. R. Petersen. 1982. *Primate communication*. Cambridge: Cambridge University Press.
- Snowdon, C. T., and S. de la Torre. 2002. Multiple environmental contexts and communication in pygmy marmosets (*Cebuella pygmaea*). *Journal of Comparative Psychology* 116: 182–88.
- Snowdon, C. T., and M. Hausberger. 1997. *Social influences on vocal development*. Cambridge: Cambridge University Press.
- Stokoe, W. C. 2001. *Language in hand: Why sign came before speech*. Washington, D.C.: Gallaudet University Press.
- Stringer, C. 2002. The morphological and behavioural origins of modern humans. In *The speciation of modern Homo sapiens*, ed. T. J. Crow, 23–30. Oxford: Oxford University Press. [MCC]
- Suddendorf, T., and M. C. Corballis 2007. The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences* 30:299–351. [MCC]
- Tanner, J. E. 1998. *Gestural communication in a group of zoo-living lowland gorillas*. Ph.D. thesis, University of St. Andrews.
- . 2004. Gestural phrases and gestural exchanges by a pair of zoo-living lowland gorillas. *Gesture* 4:25–42.
- Tanner, J. E., and R. W. Byrne. 1993. Concealing facial evidence of mood: Perspective-taking in a captive gorilla. *Primates* 34:451–57.
- . 1996. Representation of action through iconic gesture in a captive lowland gorilla. *Current Anthropology* 37: 162–73.
- . 1999. The development of spontaneous gestural communication in a group of zoo-living lowland gorillas. In *The mentalities of gorillas and orangutans: Comparative perspectives*, ed. S. T. Parker, R. W. Mitchell, and H. L. Miles, 211–39. Cambridge: Cambridge University Press.
- Tanner, J. E., F. G. Patterson, and R. W. Byrne. 2006. Development of spontaneous gestures in zoo-living gorillas and sign-taught gorillas: From action and location to object representation. *Journal of Developmental Processes* 1:69–103. [JET]
- Terrace, H., L. Petitto, R. Sanders, and T. Bever. 1979. Can an ape create a sentence? *Science* 206:891–902.

- Thorpe, S. K. S., R. L. Holder, and R. H. Crompton 2007. Origin of human bipedalism as an adaptation for locomotion on flexible branches. *Science* 316:1328–31. [MCC]
- Tomasello, M. 1995. Joint attention as social cognition. In *Joint attention: Its origin and role in development*, ed. C. Moore and P. J. Dunham, 103–30. Hillsdale, N.Y.: Lawrence Erlbaum.
- Tomasello, M., and J. Call. 1997. *Primate cognition*. New York: Oxford University Press.
- Tomasello, M., J. Call, and B. Hare. 2003. Chimpanzees understand psychological states: The question is which ones and to what extent? *Trends in Cognitive Sciences* 7:153–56.
- Tomasello, M., J. Call, K. Nagell, R. Olguin, and M. Carpenter. 1994. The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates* 35: 137–54.
- Tomasello, M., J. Call, J. Warren, T. Frost, M. Carpenter, and K. Nagell. 1997. The ontogeny of chimpanzee gestural signals. In *Evolution of communication*, ed. S. Wilcox, B. King, and L. Steels, 224–59. Amsterdam: John Benjamins.
- Tomasello, M., M. Carpenter, J. Call, T. Behne, and H. Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28:1–17.
- Tomasello, M., B. L. George, A. C. Kruger, M. J. Farrar, and A. Evans. 1985. The development of gestural communication in young chimpanzees. *Journal of Human Evolution* 14:175–86.
- Tomasello, M., D. Gust, and G. T. Frost. 1989. A longitudinal investigation of gestural communication in young chimpanzees. *Primates* 30:35–50.
- Tomasello, M., S. Savage-Rumbaugh, and A. Kruger. 1993. Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development* 64:1688–1705. [JET]
- Tomasello, M., and K. Zuberbühler. 2002. Primate vocal and gestural communication. In *The cognitive animal: Empirical and theoretical perspectives on animal cognition*, ed. M. Bekoff, C. S. Allen, and G. Burghardt, 293–99. Cambridge: MIT Press.
- Uhlenbroek, C. 1996. The structure and function of the long-distance calls given by male chimpanzees in Gombe National Park. Ph.D. thesis, University of Bristol.
- Ulbaek, I. 1998. The origin of language and cognition. In *Approaches to the evolution of language: Social and cognitive bases*, ed. J. R. Hurford, M. Studdert-Kennedy, and C. Knight, 30–43. Cambridge: Cambridge University Press. [CK]
- Van Hooff, J. A. R. A. M. 1962. Facial expressions in higher primates. *Symposium of the Zoological Society of London* 8: 7–68.
- . 1967. The facial displays of the catarrhine monkeys and apes. In *Primate ethology*, ed. D. Morris, 7–68. London: Weidenfeld and Nicolson.
- . 1971. *Aspeten van het social gedrag en de communicatie bij humane en hogere niet-humane primaten*. Ph.D. diss., University of Utrecht.
- . 1973. A structural analysis of the social behaviour of a semi-captive group of chimpanzees. In *Social communication and movement: Studies of interaction and expression in man and chimpanzee*, ed. M. von Cranach and I. Vine, 75–162. London: Academic Press.
- Vea, J. J., and J. Sabater-Pi. 1998. Spontaneous pointing behaviour in the wild pygmy chimpanzee (*Pan paniscus*). *Folia Primatologica* 69:289–90.
- Voelkl, B., and L. Huber. 2000. True imitation in marmosets. *Animal Behaviour* 60:195–202.
- . 2007. Imitation as faithful copying of a novel technique in marmoset monkeys. *PLoS One* 2:e611.
- Waters, G. S., and R. S. Fouts. 2002. Sympathetic mouth movements accompanying fine motor movements in chimpanzees (*Pan troglodytes*) with implications toward the evolution of language. *Neurological Research* 24:174–80.
- Whiten, A. 1998. Imitation of the sequential structure of actions by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 112:270–81.
- Wilkins, David. 2003. Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In *Pointing: Where language, culture, and cognition meet*, ed. Sotaro Kita, 171–215. Hillsdale, N.J.: Lawrence Erlbaum. [DAL]
- Wilson, M. L., M. D. Hauser, and R. W. Wrangham. 2001. Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour* 61:1203–16.
- Wohlschläger, A., M. Gattis, and H. Bekkering. 2003. Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society B* 358:501–15.
- Woodruff, G., and D. Premack. 1979. Intentional communication in the chimpanzee: The development of deception. *Cognition* 7:333–52.
- Zeller, A. 1980. Primate facial gestures: A study of communication. *International Journal of Human Communication* 13:565–606.
- Zuberbühler, K., ed. 1999. *Interspecific semantic communication in two forest monkeys*. Paper presented at the XXVI International Ethological Conference, August 2–9, 1999, Bangalore, India.
- . 2000a. Interspecific semantic communication in two forest monkeys. *Proceedings of the Royal Society B* 267: 713–18.
- . 2000b. Referential labelling in Diana monkeys. *Animal Behaviour* 59:917–27.
- . 2001. Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behavioral Ecology and Sociobiology* 50:414–22.
- . 2002. A syntactic rule in forest monkey communication. *Animal Behaviour* 63:293–99.