

Constraints and preadaptations in the earliest stages of language evolution

DOROTHY L. CHENEY AND ROBERT M. SEYFARTH

Abstract

If we accept the view that language first evolved from the conceptual structure of our pre-linguistic ancestors, several questions arise, including: What kind of structure? Concepts about what? Here we review research on the vocal communication and cognition of nonhuman primates, focusing on results that may be relevant to the earliest stages of language evolution. From these data we conclude, first, that nonhuman primates' inability to represent the mental states of others makes their communication fundamentally different from human language. Second, while nonhuman primates' production of vocalizations is highly constrained, their ability to extract complex information from sounds is not. Upon hearing vocalizations, listeners acquire information about their social companions that is referential, discretely coded, hierarchically structured, rule-governed, and propositional. We therefore suggest that, in the earliest stages of language evolution, communication had a formal structure that grew out of its speakers' knowledge of social relations.

That which distinguishes man from the lower animals is not the understanding of articulate sounds, for, as every one knows, dogs understand many words and sentences . . . It is not the mere articulation which is our distinguishing character, for parrots and other birds possess this power. Nor is it the mere capacity of connecting definite sounds with definite ideas; for it is certain that some parrots, which have been taught to speak, connect unerringly words with things, and persons with events. The lower animals differ from man solely in his almost infinitely larger power of associating together the most diversified sounds and ideas; and this obviously depends on the high development of his mental powers. (Charles Darwin, 1871, *The Descent of Man*)

1. Introduction

Since at least the time of Locke (1690), a central tenet of the philosophy of language has held that language derives from our experience with objects, events, and the relations we perceive in the world. This assumption underlies many studies of children's language learning, where it is generally accepted that "a well-articulated conceptual structure predates language and is (one of) its required causal engines" (Fisher and Gleitman 2003). It also appears in theories of language evolution, where several authors have suggested that "the roots of grammar lie in prehuman conceptual structure" (Newmeyer 2003; see also Pinker and Bloom 1990; Jackendoff 1987, 2002; Kirby 1998; Newmeyer 1991; Hurford 1998, 2003).

This hypothesis, however, begs several questions: What kind of structure? Concepts about what? To explain the earliest stages of language evolution, one would ideally like to describe communication and cognition in the common ancestor of human and nonhuman primates and then identify the first significant changes that might logically have led, over the next several million years, to modern human language.

Few discussions of language evolution, however, start with this common ancestor. Indeed, most speculations about the evolution of language make only passing reference to the natural communication and cognition of non-linguistic primates, choosing instead to emphasize the most complex properties of modern languages (Pinker and Bloom 1990; Pinker 1994; Deacon 1997; Jackendoff 1987, 2002).

Others have searched for the origins of language in the tools and cave paintings left by early members of the genus *Homo* (e.g., Mithen 1996), through theoretical models (e.g., Nowak et al. 2000; Nowak and Komarova 2001), or

by drawing analogies between proto-language and the simple pidgins used by some modern humans (Bickerton 1990; Calvin and Bickerton 2000). Either implicitly or explicitly, however, all such attempts have concerned a hypothetical creature that already had the capacity to represent abstract thoughts symbolically (in the case of artifacts), to imitate sounds and generate new words (in the case of many models), or to use a simple grammar (in the case of pidgins). In other words, they have concerned a species that likely had long since diverged from the common ape/human ancestor and already had many of the cognitive and communicative abilities that we regard as uniquely human. Conversely, attempts to infer language abilities by reconstructing the anatomy of early human vocal tracts (e.g., Lieberman 1984, 1991) have not considered the cognitive abilities that would have made a human vocal tract adaptive.

Given the extraordinary complexity of language and the daunting problems it poses for language learners, it is easy to see why so many linguists have concluded that to explain language evolution one must explain the evolution of syntax and semantics in all of their complexity. This is not, however, where an account of language evolution ought to begin. Fully human syntax arguably emerged relatively late in evolution, after a period of time when other, simpler features of language and cognition had already appeared. These simpler features would have been adaptive in their own right and, more importantly, would have constituted a new system of communication that, unlike its predecessors, created the preconditions needed to favor the evolution of more complex features like syntax.

What were these simpler features? In the sections below we review some recent research on the vocal communication and cognition of nonhuman primates, focusing in particular on results that may be relevant to the earliest stages of language evolution. From these data we draw several conclusions. First, nonhuman primates' inability to represent the mental states of others makes their communication fundamentally different from human language. Second, while nonhuman primates' production of vocalizations is highly constrained, their ability to extract complex information from sounds is not. Upon hearing vocalizations, listeners acquire information that is referential, highly structured, propositional, and rule-governed. Third, at least some of the components of prehuman cognitive structure that may have given rise to the evolution of language are evident in primates' knowledge of social relationships.

2. Theory of mind and the use of vocalizations by non-human primates

Baboons (genus *Papio*) are Old World monkeys that shared a common ancestor with humans approximately 36 million years ago (Boyd and Silk 2000). They live in groups of 50–150 throughout the savannah woodlands of Africa. Al-

though most males emigrate to other groups as young adults, females remain in their natal groups throughout their lives, maintaining close social bonds with their matrilineal kin. Females can be ranked in stable, linear dominance hierarchies that define priority of access to scarce resources. Offspring acquire ranks similar to those of their mothers. The stable core of a baboon group, therefore, is a hierarchy of matrilines (for reviews see Melnick and Pearl 1987; Silk et al. 1999).

When moving through wooded areas of their range, baboons often give loud *contact* barks (Cheney et al. 1996; Rendall et al. 2000; Fischer et al. 2000). Because barks from widely separated individuals are typically temporally clumped, baboons often appear to be answering one another. It is unclear, however, whether individuals give barks with the intent of informing others of their location, or whether these calls simply reflect a signaler's own state of separation from the group. These two explanations are functionally equivalent, because in either case listeners could use contact barks to determine the location of other group members. However, they are based on very different underlying mental mechanisms. The first explanation assumes that callers recognize that other individuals may be ignorant of the group's location even when they are not. It assumes that baboons have a 'theory of mind' (Premack and Woodruff 1978). The second makes no such assumptions. It requires only that callers vocalize when they are anxious, and that listeners learn that contact barks inform them of the location of at least one other individual.

The distinction between these two explanations is crucially important to theories of language evolution. If, as Grice (1957) and others have argued, true linguistic communication cannot occur unless both speaker and listener take into account each other's state of mind, then monkeys cannot be said to communicate unless they use calls like contact barks with the intent to provide information to others. By contrast, if monkeys cannot recognize the distinction between their own knowledge and somebody else's, calls that serve an informative function will be based on mental mechanisms that differ fundamentally from those found in human speech. Hypotheses based on mental state attribution predict that individuals will answer the contact barks of others even when they themselves are in the center of the group progression and at no risk of becoming lost. Alternatively, if baboons are incapable of understanding that other individuals' mental states can be different from their own, they should be unable to recognize when another individual thinks that she has become lost unless they themselves face a similar risk. Under these circumstances, contact barks will simply reflect the state and location of the signaler.

Playback experiments support the latter interpretation. When females were played a relative's contact bark from a distant loudspeaker, they gave an 'answering' bark only when they themselves were separated from the group (Cheney et al. 1996). This was true even of contact barks given by their own

infants (Rendall et al. 2000). (By analogy, imagine the response of a human parent in a busy store who hears her child crying that she is lost.) Calling appears to be an emotional reaction to a stressful and potentially dangerous situation. Although most contact barks elicit no answering calls, giving a bark potentially benefits the signaler because it prompts listeners to assess their own positions (however reflexively) and perhaps give a bark themselves, thereby alerting the signaler to the location of at least one individual. Because barks (like many primate calls) are individually distinctive (reviewed by Snowdon 1990, Ghaz and Santos 2004), and because there is a close link between the eliciting stimulus and the production of a specific type of call, listeners recognize immediately both the signaler's identity and her location. Listeners thus acquire information from a signaler who did not, in the human sense, intend to provide it.

The lack of a theory of mind constitutes a fundamental difference between human language and nonhuman primate communication. In humans, perception of another individual's mental state is perhaps the most common stimulus eliciting speech: we talk to inform, persuade, or in some way change what another thinks. By contrast, while monkeys may recognize the effect that their vocalizations have on their audience's behavior, they seem largely unaware of their effect on their audience's knowledge and beliefs (Cheney and Seyfarth 1990, 1997b; Seyfarth and Cheney 2003b).

Monkeys clearly lack language, so it is not surprising that they lack a full-blown theory of mind. This is not to say, however, that organisms lacking language are entirely incapable of inferring intent and motives in others. In children, inferences about other individuals' attention, motives, and directedness constitute an early precursor to language learning and full mental state attribution (see, e.g., Malle et al. 2001). Long before they are able to attribute false beliefs to others, children can recognize the focus of another individual's intentions and other individuals' likes and dislikes (see below).

Similarly, despite their apparent inability to distinguish other animals' mental states from their own (e.g., Povinelli and Eddy 1996; reviewed in Tomasello and Call 1997), chimpanzees appear to have a rudimentary understanding of the relation between seeing and knowing. Experiments conducted by Hare et al. (2000, 2001) suggest that, in at least some situations, chimpanzees know what a conspecific has or has not seen, and from this information may infer what a conspecific does or does not know. In contrast, monkeys seem unable to recognize the referential nature of attention. In the absence of training, for example, they are unable to use pointing or gaze direction to locate hidden objects (Anderson et al. 1996; Tomasello and Call 1997; Vick et al. 2001).

Nonetheless, monkeys are very attentive to other individuals' gaze, particularly when it is directed at themselves (Perrett and Emery, 1994; Emery, 2000). They use gaze to target opponents and to recruit other individuals' support in aggressive alliances (Sato and Nakamura 2001; Tomasello et al. 2001). At a

rudimentary level, then, monkeys may use the information obtained from another individual's direction of gaze to make inferences about her intentions – to infer, for example, that a threat gesture is directed at themselves and not at someone else.

Social interactions among baboons are mediated not only by different behaviors associated with affiliation and aggression, but also by a variety of vocalizations. Their most common vocalization is a low amplitude grunt, which appears to signal benign intent and facilitates social interaction (Cheney et al. 1995b; Silk et al. 1996). Because baboons live in large social groups, however, the intended recipient of a grunt is not always immediately apparent. For example, in a typical social interaction a dominant female approaches two subordinate females, one of whom has an infant, and utters a grunt. The female with the infant remains seated, while the other moves away. How does a female decide whether to stay or leave? To a large degree, her response may depend upon gaze direction and relatively simple learned contingencies. Vocalizing females often appear to be looking at one particular individual, and recipients are doubtless sensitive to gaze direction. Similarly, both the mother and the female without an infant may have learned that grunts are correlated with the handling of an infant. As a result, the female without an infant moves away while the mother remains seated (Cheney et al. 1995b; Silk 2002).

When accompanying visual signals are absent, however, the interpretation of vocal signals may be more difficult. Baboons often forage in woodland where other individuals can be heard but not seen. Vocal signals like grunts are individually distinctive (Owren et al. 1997; Rendall et al. 1999), so listeners can readily ascertain an unseen caller's identity. In the absence of visual cues, however, listeners may find it difficult to determine the caller's intended recipient, and hence the appropriate behavioral response. Again, relatively simple contingencies, including the nature of recent interactions, may guide listeners' responses.

For example, after 13 % of aggressive disputes among adult female baboons, dominant females grunt to their former victims. These grunts appear to serve a reconciliatory function, both reducing the anxiety of subordinate females (Cheney et al. 1995) and influencing their subsequent interactions. Cheney and Seyfarth (1997a) used a playback experiment to mimic vocal reconciliation. Subordinate subjects were played the grunt of their former opponent in the minutes immediately following a dispute and then observed for half an hour. As controls, subjects heard either no grunt at all or the grunt of another dominant female who had not been involved in the earlier dispute. After playbacks of reconciliatory grunts, subjects approached their former opponents and tolerated their opponents' approaches at significantly higher rates than they did in the absence of a grunt. In contrast, when subjects heard either a control grunt from another female or no grunt, they avoided their former opponents. Sig-

nificantly, subjects' interactions with control females were unaffected by call playbacks. Simply hearing a more dominant female's grunt did not increase the likelihood that subjects would approach that female.

These results are relevant to hypotheses concerned with monkeys' ability to judge the directedness of vocal signals. By approaching an opponent only if they had recently heard her reconciliatory grunt, subjects behaved as if they interpreted the grunt as a signal of the opponent's friendly intent toward them. In contrast, by failing to respond to the grunt of an uninvolved female, they acted as if they interpreted that call as directed at someone else.

Subordinate subjects' responses were not simply the result of priming due to a recent interaction. In subsequent experiments, Engh et al. (in press) played a dominant female's threat-grunts to a subordinate subject shortly after that female had either groomed or fought with the subject. Subjects responded significantly more strongly to the threat-grunt after recent aggression. Moreover, when the threat-grunt playback had been preceded by aggression rather than grooming, subjects were significantly less likely to approach the dominant signaler within the next 15 minutes and less likely to tolerate the dominant signaler's approach.

Taken together, these experiments suggest that female baboons are able to make inferences about the directedness of vocalizations even in the absence of visual cues, and that the nature of prior interactions affects their subsequent behavior. Upon hearing a vocalization, listeners need to predict what behavior is likely to follow. In doing so, they must take into account the identity of the signaler, call type (is it associated with aggression?), the nature of their prior interactions with the signaler (were they aggressive, friendly, or neutral?), and the correlation between past interactions and future ones (does a recent grooming interaction lower or increase the likelihood of aggression?). No doubt contingency learning plays a major role in each of these assessments. However, because listeners' responses depend on simultaneous consideration of all of these factors, however, this learning is likely to be both complex and subtle.

Moreover, explanations based on behavioral contingencies cannot explain some aspects of listeners' behavior, in particular their ability to predict whether a call is directed at themselves or at some other individual. For example, in the earlier reconciliation experiments (Cheney and Seyfarth 1997a), subjects who heard their opponent's reconciliatory grunt following a fight were even more likely to approach their opponent than they were under baseline conditions. If listeners' responses were guided only by learned contingencies, they should have associated the reconciliatory grunt simply with a low probability of aggression. It should have returned their behavior to baseline tolerance levels, but it should not have induced them to increase their interactions with the signaler. It is difficult to explain subjects' behavior except by arguing that females interpreted the grunts as targeted specifically at themselves, as a directed signal

of benign intent. We may speculate that monkeys have a rudimentary recognition of other individuals' intentions toward themselves. If true, this would represent a crucial first step toward a communication system like language, in which speakers and listeners routinely take into account each other's motives and knowledge.

3. The asymmetry between vocal production and comprehension

As many have pointed out, physiological constraints place an upper limit on the variety of calls that non-human primates (and other mammals) can produce. We will not belabor this point here, except to reiterate that such constraints limit not only the number and variety of sounds produced but also the ability of non-human primates to modify and invent new sounds (see Fitch and Hauser 1995; Fitch 2004; Fischer et al. 2002 for review). Our primary concerns in this essay are not the physiological but, rather, the cognitive constraints on vocal production.

While the number of distinct calls that animals produce is highly constrained, the number of sounds or signs that a parrot, dolphin, sea lion, or chimpanzee can learn to associate with a given stimulus or outcome is, if not limitless, certainly in the tens to hundreds. And while we can quibble about the exact nature of the underlying mental representation that is evoked when a dog hears *walk*, *car*, *supper*, or *ball*, the fact that these calls evoke different responses suggests strongly that they are assigned functionally distinct 'meanings' (Thompson 1995; Cheney and Seyfarth 1997b; Schusterman and Kastak 1998; Kaminski et al. 2004).

For Darwin, the intuition that animals can effortlessly learn to associate myriad sounds and signals with 'general ideas or concepts' (1871: 83) was more significant than differences in 'mere articulation.' We would argue the point slightly differently: physiological constraints on production are not just important in their own regard; equally striking is the difference between relatively constrained production and more open-ended perception. Animals' limited vocal repertoires are particularly puzzling because they appear to have so many concepts that could, in principle, be articulated.

3.1. Vocal development and modification

In marked contrast to children, who learn both to produce and to comprehend literally thousands of new words during their first three years of life, monkeys and apes rarely modify their vocal repertoires by adding new sounds. Although some primates make subtle modifications in their vocalizations as a result of experience (Hauser 1989, 1992; Elowson and Snowdon 1994; Mitani and Brandt

1994; Seyfarth and Cheney 1997; Crockford et al. 2004), a baboon in Kenya produces more or less the same sounds in the same contexts as a baboon in Botswana. This conclusion follows not only from research on many different species but also from experiments in which infant macaques were cross-fostered and raised in groups of another species (Owren et al. 1993). Subjects in these experiments showed virtually no modification in their use of vocalizations, even in social contexts where the members of their adopted species used calls that differed from their own.

This is not to say that primate vocalizations are just involuntary and reflexive. In the laboratory, some features of call production by rhesus macaques can be brought under operant control (Peirce 1985). In the wild, monkeys vary their vocal production depending upon context and the identity of social partners (e.g., Cheney and Seyfarth 1997a). The acoustically different alarm calls given by primates to different predators, though delivered in emotionally charged situations, are also not as involuntary as they first appear. In the seconds after spotting a predator, a monkey must decide whether to remain silent or give an alarm call, and, if an alarm is to be given, which call is appropriate (Seyfarth et al. 1980; Zuberbuhler et al. 1999; Fischer et al. 2001a, 2001b). The production of alarm calls by monkeys, like call production in many animals, is affected by an audience; vervet monkeys (*Cercopithecus aethiops*), for example, seldom alarm-call when they are alone and are more likely to give alarm calls in the presence of kin than non-kin (Cheney and Seyfarth 1990; see also Gyger et al. 1986 for data on chickens).

3.2. Repertoire size and vocal production

Although the vocal repertoires of monkeys and apes are undoubtedly smaller than the vocabularies of humans, it would be wrong to conclude that they consist simply of the limited number of grunts, screams, and barks that are acoustically salient to human listeners. In fact, most nonhuman primates possess an acoustically graded vocal repertoire. Despite this intergradation, acoustically different calls are often associated with qualitatively different contexts and, like humans, primates appear to perceive categorical boundaries within this acoustic continuum. They do so in two respects. First, they distinguish the calls of other group members despite the apparent lack of discrete acoustical boundaries between different individuals' voices (e.g., Snowdon 1987; Mitani 1996; Rendall et al. 1996; Cheney and Seyfarth 1997b; Rendall et al. 1999). Second, listeners recognize distinct subtypes within broad acoustic classes, each of which conveys different information. Baboons, for example, respond differently to *social* and *move* grunts (Rendall et al. 1999), to *contact* and *alarm* barks (Fischer et al. 2001a, b, 2002), and to *contest* and *alarm* wahoos (Kitchen

et al. 2003). Their different responses suggest that they assign functionally referential 'meaning' to these acoustically graded sounds (Marler et al. 1992; Hauser 1996; Rendall 2003).

The perception of call subtypes within a graded acoustic continuum increases effective repertoire size. So does call combination. Indeed, two recent studies suggest that monkeys and apes may effectively increase their vocal repertoire by combining existing calls and assigning these combinations to new contexts.

Like many forest monkeys, Campbell's monkeys (*Cercopithecus campbelli*) give acoustically different alarm calls to leopards and eagles. In less dangerous contexts, they emit a low, resounding 'boom' call prior to the alarm calls. Sympatric diana monkeys (*C. diana*) respond strongly to the Campbell's monkey alarm calls. They also appear to be sensitive to the semantic changes caused by call combination, because they no longer respond to Campbell's monkeys alarm calls if they are preceded by a boom (Zuberbuhler 2002; see also Robinson 1984; Snowdon 1990). Similarly, chimpanzees frequently combine different call types when vocalizing, and in some cases also supplement calls by drumming their hands and feet against resonant tree buttresses (Mitani 1993). In the Ivory Coast, male chimpanzees produce three acoustically different subtypes of barks: one when hunting, one when they encounter snakes, and a third, more generic bark type in a variety of different contexts. In two very limited circumstances, when traveling or encountering a neighboring group, the chimpanzees combine a bark with drumming (Crockford and Boesch 2003). This signal combination has the potential to convey information that is qualitatively different from (and more specific than) the information conveyed by a single call type.

Depending upon the definition one chooses, these call combinations may qualify as syntactical. Marler (1977), for example, distinguished between phonological syntax, in which call combinations carry a meaning that is more than just the sum of their parts, and lexical syntax, in which the component parts also play functional roles as subjects, verbs, modifiers, and so on. According to this distinction, the call combinations discussed above may be examples of phonological, but perhaps not lexical, syntax (but see Zuberbuhler 2002 for a slightly different view).

3.3. *The mechanisms underlying vocal production*

What mental representations, if any, underlie vocal production? When a female baboon gives a contact bark, does she carry some concept of separation that is different from the representation that is instantiated when she gives a similar-sounding alarm bark? The fact that each bark subtype is given in a relatively

specific context seems to support the hypothesis that each call is an expression of a specific mental concept. Alternatively, the fact that female baboons throughout Africa give harsh, noisy barks when encountering predators and tonal barks when separated from the group suggests that links between sound and concept are far more constrained in baboons than in human languages.

Why might acoustically similar calls serve such markedly different functions? When baboons hear an alarm bark they flee into trees; when they are separated from the group and hear an acoustically similar contact bark, they descend from trees and run to join their companions. Given these very different responses, it seems that selection should have favored contact and alarm calls that sound unambiguously different, rather than ones that sound alike. Their acoustic similarity may arise because they are both manifestations of similar underlying affective states (Owren and Rendall 1997). Predator encounters and group separation are both potentially dangerous contexts associated with high states of arousal. The emotions they elicit may be similar enough to lead to similar-sounding calls.

If this affect-based hypothesis is correct, however, it must also account for the subtle but consistent acoustic differences between contact and alarm barks – presumably by suggesting that a caller's internal states in these two contexts, while broadly similar, are nonetheless consistently and subtly different. It also requires that all signalers 'agree' about which affective state goes with which context. Moreover, if calls are manifestations of specific affective states that are tied to specific contexts, how do we explain signal combinations such as those used by chimpanzees when drumming? Snake barks alone are given to snakes; combined with drumming, they are given to neighboring groups. Without resorting to a convoluted argument that neighbors somehow evoke emotions that are both snake-related and aggressive (hence the drumming), it is hard to explain this difference solely by positing a shift in affective state. Just as humans assign similar-sounding morphemes or morpheme combinations to distinctly different referents, chimpanzees may have in mind two distinct concepts, one that is expressed by a bark alone and the other by a bark plus drumming. This, of course, is simply an hypothesis: the mental mechanisms underlying primate vocal production remain a largely unexplored puzzle.

3.4. The perception of 'referential' vocalizations

Compared with call production, call perception in nonhuman primates exhibits many more parallels with human speech. Listeners appear to be very sensitive to the contingencies that surround call production – contingencies that allow them to derive specific information about both caller identity and context. In captive settings, primates readily learn the voices of new caretakers and the

'meaning' of other novel auditory stimuli (like the jangling of keys) that would not ordinarily be part of their natural environment. As Darwin noted, animals seem to have an almost limitless ability to learn new sound-meaning pairings.

While call production shows little modification during development, call perception is readily modified through experience. In the macaque cross-fostering experiments mentioned above, call perception was clearly more labile than call production. Whereas cross-fostered individuals showed almost no change in vocal production, both they and the animals in their adoptive groups learned to recognize and respond appropriately to each others' vocalizations (Seyfarth and Cheney 1997). Similarly, young vervet monkeys must learn to distinguish and respond appropriately to the different alarm calls given by both their own species and nearby ungulates and birds (Seyfarth and Cheney 1990; Hauser 1988). Probably because they cannot recognize ignorance in others, adults do not actively instruct their offspring in this process. Instead, infant and juvenile primates learn to distinguish which species are dangerous and which calls function as alarms through associative processes (Cheney and Seyfarth 1990; Fischer et al. 2000).

Many primate vocalizations exhibit two properties that allow listeners to extract specific information from them. First, some calls have a high "informative value" (Seyfarth and Cheney 2003b). If a given call type is elicited only by eagles and eagles rarely appear without eliciting that call, then that call has potential to provide listeners with reliable information about the presence of an eagle (Rescorla 1988). When adult vervet and diana monkeys detect an eagle or hear an eagle's shriek, they produce an acoustically distinct alarm call; they do not produce this call when eagles are absent (Cheney and Seyfarth 1990; Zuberbuhler et al. 1999). Eagle alarm calls are thus highly informative.

Second, some calls with a high informative value also have a high "referential specificity" (Seyfarth and Cheney 2003b). For example, diana monkey *eagle* alarms are predictive not only about the presence of an eagle but also about the presence of a specific species of eagle, because the monkeys give eagle alarms only to the crowned eagle (Zuberbuhler et al. 1997, 1999). By contrast, diana monkeys give an *alert* call to a wide variety of stimuli, including mammalian and avian predators, large non-predatory animals, falling trees, and social disturbances within the group (Zuberbuhler et al. 1997). Alert calls are vaguely informative (listeners learn that *something* is happening), but they have a low referential specificity.

In habituation-dishabituation experiments, diana monkey females who hear a male's *leopard* alarm call respond by giving leopard alarm calls of their own. If they then hear a leopard's growl they show little or no response, apparently because the growl is now redundant. They do give alarm calls, however, if they are played the shriek of an eagle (they give eagle alarm calls). Conversely, females who first hear a male diana's eagle alarm call do not subsequently

respond to the shriek of an eagle, but do respond with leopard alarms if they hear the growl of a leopard (Zuberbuhler et al. 1999). Diana monkey leopard alarms and leopard growls are very different acoustically, as are diana monkey eagle alarms and eagle shrieks. Nonetheless, the monkeys treat the two leopard-associated noises and the two eagle-associated noises as if they provide the same information.

One interpretation of these results argues that the mechanisms underlying call production and perception include the formation of some mental representation about the eliciting stimulus, and this information determines whether a call will be produced and how it will be interpreted (Zuberbuhler et al. 1999). Such an interpretation suggests parallels between the processing of calls by monkeys and humans. Just as humans respond to words by noting both their acoustic properties (auditory processing) and their meaning (semantic processing), monkeys process calls at two levels and can, in some circumstances, group acoustically different calls together based on shared meaning (Seyfarth and Cheney 2003b; Zuberbuhler 2003). The meaning of these ‘words’ is defined not just by the relation between words and objects but by the relation between one word and another (cf. Deacon 1997; Pinker and Jackendoff 2005).

Our point is not to argue that these vocalizations function exactly like human words, but to suggest that nonhuman primate vocalizations exhibit just the sort of rudimentary representational properties that we would expect to find in the communication of a pre-linguistic human ancestor. Natural selection has favored the evolution of calls that convey highly specific information and the evolution of cognitive abilities in listeners that allow them to assess and compare signals according to their meaning. Diana monkeys confront a world in which there are statistical regularities; from these they learn to place a leopard’s growl and one of their own species’ alarm calls into the same functional class. Primate calls acquire their referential features at least in part because selection has favored listeners who acquire as much information as possible from the auditory stimuli that they encounter.

3.5. Vocalizations and social knowledge

The rich information acquired by listeners from relatively impoverished signals is nowhere more evident than in monkeys’ responses to vocalizations given during social interactions. Many of these calls follow highly predictable patterns. Female baboon threat-grunts, for example, are given only by higher-ranking individuals to those lower-ranking in the dominance hierarchy, while screams are given only by lower-ranking individuals to those of higher status. By attending to such calls, monkeys can deduce not only the identities of participants in an unseen interaction but also the nature of the interaction, its

direction, and its outcome.

For example, in an experiment designed to test baboons' knowledge of other individuals' dominance ranks, subjects were played a sequence of calls that mimicked a rank reversal among two unrelated females (Cheney et al. 1995b). Call sequences consisted of a series of grunts combined with a series of *fear* barks. As mentioned earlier, grunts are individually distinctive calls that function as signals of benign intent; they are given both by high-ranking females to lower-ranking ones and vice versa. Fear barks, however, are unambiguous signals of subordination. They are always given by subordinate females to more dominant ones, often when the more dominant is grunting and attempting to interact with the subordinate's infant.

Subjects were played a causally inconsistent call sequence in which a low-ranking female apparently grunted to a high-ranking female and the high-ranking female gave fear barks. As a control, subjects heard the same sequence of grunts and fear barks made causally consistent by the inclusion of additional grunts from a third female who was dominant to both of the others. Regardless of their own relative ranks, subjects responded by looking toward the speaker for significantly longer durations to inconsistent sequences that violated the female dominance hierarchy. They appeared to recognize not only the identities of different signalers but also the rank relations that existed among them (Cheney et al. 1995a; Cheney and Seyfarth 2004).

Other playback experiments demonstrate that baboons can integrate their knowledge of other individuals' dominance ranks and kinship to recognize that their social group is composed of a hierarchy of families (Bergman et al. 2003; for evidence of baboons' recognition of other individuals' kinship relationships, see Cheney and Seyfarth 1999). Female subjects were played sequences of threat-grunts and screams that mimicked a fight between two other females. One sequence consisted of an anomalous threat-grunt-scream sequence mimicking a *within-family* rank reversal, for example B₃ threat-grunts and B₁ screams (in this notation, letters are used to denote matrilineal families, ranked A, B, C, etc. from top to bottom and numbers are used to denote individuals within these families). A second consisted of an anomalous sequence mimicking a *between-family* rank reversal (e.g., C₁ threat-grunts + B₃ screams). The third consisted of a no-reversal control sequence consistent with the current female dominance hierarchy (e.g., B₃ threat-grunts and C₁ screams, or B₁ threat-grunts and B₃ screams).

If baboons classify other individuals simultaneously according to both their individual attributes (rank) and their membership in a higher-order class (matrilineal kin group), they should have responded most strongly to the apparent between-family rank reversal than to the within-family rank reversal. A between-family rank reversal is potentially much more significant than a within-family rank reversal because it signals a possible change in the dominance

relations of two entire matrilineages rather than just two individuals within the same family. This was indeed the case. In contrast, subjects responded relatively weakly both to sequences that mimicked a within-family rank reversal and to those in which no reversal took place (Bergman et al. 2003). Below we consider the implication of these results for theories of language evolution.

3.6. Summary: The asymmetry between caller and recipient

In nonhuman primates – and perhaps many other animals – vocal production and perception are strikingly different. Where call production is concerned, primates have a comparatively small repertoire of vocalizations whose acoustic features are relatively innate. And while vocal production is not entirely involuntary, most call types are tightly linked to a specific stimulus or context.

In contrast, primates and other animals exhibit an almost open-ended ability to learn novel sound-meaning pairs. Learning can occur throughout adulthood and is not linked to a particular social context. Moreover, the information acquired from vocalizations is organized in highly structured ways and may be both representational and propositional (see below).

The discontinuities between production and perception result in an oddly unbalanced form of communication: monkeys (and other animals) can learn many sound-meaning pairs but cannot produce new words, and they understand conceptual relations but cannot attach labels to them (Cheney and Seyfarth 1997b). At the same time, it is crucial to remember that listeners are also signalers. The baboon who artfully decodes an anomalous threat-grunt-scream sequence into a rank reversal in the female dominance hierarchy will on some other occasion be the individual who is screaming. It therefore seems unlikely that the mental representations that underlie the interpretation and classification of calls are qualitatively different from those that accompany call production, or – more often – those that remain private and unexpressed. The natural communication of nonhuman primates argues strongly against the view that a concept cannot be acquired unless it is instantiated in one's language (reviewed by Gleitman and Papafragou 2005). Monkeys and apes have many concepts for which they have no words.

4. Pulling it all together

4.1. The crucial importance of a theory of mind

Why should an animal that can learn to associate hundreds of sounds and symbols with objects and events find it so difficult to produce novel calls or create

novel call combinations? To answer this question it may be useful to contrast word learning in animals with word learning in very young children.

By the age of one year, children already seem to understand that words can be mapped onto objects and actions in the world (reviewed by Fisher and Gleitman 2002; Bloom 2003). Crucially, this understanding seems to be accompanied by a form of 'social referencing', in which the child uses other people's direction of gaze, gestures, and emotions to appraise a situation. Children at this age also actively attend to the speaker's gaze and focus of attention when inferring the referent of the speaker's utterance, as if they have developed some tacit understanding that gaze and attention are a reflection of underlying knowledge (Baldwin 1993). Similarly, around the age of one year children begin to use gestures and sounds to recruit adults' attention to external objects or events. Through imitation, declarative gestures, and speech, they demonstrate their view of adults as intentional beings. Children's ability to compare another's perceptual state with their own forms the basis of a social referencing system that is integral to early word learning (Bloom and Markson 1998; Tomasello 2003).

Although there are precursors to these abilities in the social interactions and communication of monkeys and apes, they remain rudimentary (e.g., Cheney and Seyfarth 1990; Anderson et al. 1996; Tomasello and Call 1997). Baboons recognize when calls are being directed at themselves and they seem to have some understanding of other individuals' intentions (Cheney and Seyfarth 1997a; Engh et al. in press). In contrast to the communication of even very young children, however, monkey vocalizations appear designed to influence other individual's behavior, not their attention or knowledge. Although monkeys vary their calling rates depending upon the presence and composition of their audience, they do not act deliberately to inform ignorant individuals, nor do they attempt to correct or rectify false beliefs in others or instruct others in the correct usage or response to calls (Seyfarth and Cheney 1986).

The inability of monkeys and perhaps also apes to recognize many of the mental mechanisms that underlie communication may also partially explain the apparent absence of lexical syntax in their vocalizations. At least some of the sounds produced by monkeys are functionally referential (Cheney and Seyfarth 1990; Zuberbuhler 2003). Because they lack syntactic properties, however, their exact meaning is imprecise. A vervet monkey's leopard alarm, for example, cannot really be described as a command to action ("*Run into the trees*") because not all vervets run into trees upon hearing the call, and vervets already in trees will also give this call if they spot a leopard. Nor is the call simply a noun ("*Leopard*" or "*Carnivore*"), because it consistently evokes a flight response from at least some listeners. Instead, the vervet's leopard alarm call seems best described as a proposition: a single utterance or thought that simultaneously incorporates a subject and a predicate.

Vervet alarm calls are thus simultaneously both eventish and objectish, because they incorporate both reference to an object and a disposition to behave toward that object in a particular way. They refer to a particular sort of immediate danger, and they function to designate particular classes of predators. There is no evidence, though, that a leopard alarm call can be modified to elaborate upon the characteristics of the leopard in question. Through repetition and changes in amplitude (both of which seem to carry prosodic information), alarm calls may inform others of the immediacy of danger (e.g., Manser et al. 2001). They cannot, however, specify whether a leopard is big or small, sleeping or stalking, in a tree or on the ground.

The lack of syntax in nonhuman primate vocalizations cannot be traced to an inability to recognize argument structure – to understand that an event can be described as a sequence in which an agent performs some action on an object. Baboons clearly understand the difference between *Hannah threatens Sylvia* and *Sylvia threatens Hannah*. It is also likely that nonhuman primates can represent descriptive modifiers (a *big* leopard as opposed to a *small* one) and prepositions that specify locations (a leopard *in* a tree as opposed to one *under* the tree). Captive dolphins (Herman et al. 1993), sea lions (Schusterman and Krieger 1986), and African grey parrots (Pepperberg 1992) can be taught to understand and in some cases even produce modifiers. In their natural behavior, therefore, nonhuman primates certainly act as if they are capable of thinking (as it were) in sentences. But even if monkeys mentally tag events as argument structures (who does what to whom), they fail to map these tags onto a communicative system in any stable or predictable way.

In sum, the communication of non-human animals lacks three features that are abundantly present in the utterances of young children: a rudimentary ability to attribute mental states different from their own to others, the ability to generate new words, and lexical syntax. We suggest that the absence of all three features is not accidental, and that the lack of one (a theory of mind) may explain the lack of the others (words and syntax). Because they cannot attribute mental states like ignorance to one another and are unaware of the causal relation between behavior and beliefs, monkeys and perhaps also apes do not actively seek to explain or elaborate upon their thoughts. As a result, they are largely incapable of inventing new words and of recognizing when thoughts should be articulated. Because monkeys cannot distinguish between what they know and what others know, they fail to recognize that ignorant individuals must have events explained to them. They do not understand the need to specify whether a leopard is *in* a tree or *on* the ground, nor do they comment about things in their absence (“*the flea bag that ate my mother*”). Instead, monkeys’ calls reflect the knowledge the signaler has rather than the knowledge the signaler intends his audience to acquire (Cheney and Seyfarth 1990, 1996).

A thoughtful reader may object to at least several parts of this proposal. What do we mean by stating that a rudimentary theory of mind is essential to word learning and production? After all, birds learn songs, but few would ascribe song learning in birds to a theory of mind. Compared to birds, however, the vocabularies of humans are more open-ended, more context-independent, and more functionally eclectic. New sounds can be assigned to almost any object, event, or descriptor. Equally important, the adoption of new words within the community is rapid and reciprocal.

Crucially, even very young children appear to have some tacit recognition of this relationship between comprehension and production. As a result, they understand that they can use a newly acquired word to influence another's behavior or attention. Similar capacities have yet to be demonstrated in the natural communication of any non-human primate.

4.2. *Social intelligence as a precursor of language*

Despite their lack of a theory of mind, nonhuman primates' knowledge of social relationships and call meaning exhibits several striking properties.

First, knowledge is *representational*: when a monkey hears a vocalization she acquires information that is very specific – about a particular sort of predator or a particular kind of interaction between specific individuals.

Second, social knowledge is based on properties that have *discrete values*: for example, individual identity, sex, and dominance rank (Worden 1998).

Third, animals combine these discrete-valued traits to create a representation of social relations that is *hierarchically structured*: baboons, for example, create a nested hierarchy in which others are placed in a linear rank order and simultaneously grouped according to matrilineal kinship in a manner that preserves ranks both within and across families.

Fourth, social knowledge is *rule governed and open-ended*: baboons recognize that vocalizations follow certain rules of directionality (for example, screams are only given by subordinates to dominants), and that directionality should conform to the existing dominance hierarchy. The hierarchy is open-ended because new individuals can be added or eliminated without altering the underlying structure, and because the set of all possible interactions is very large (Worden 1998; Seyfarth and Cheney 2003a).

Fifth, knowledge is *propositional*: In tests mimicking between-family rank reversals, listeners responded as if they parsed a call sequence as a dramatic narrative: “Hannah is threatening Sylvia and Sylvia is screaming. But Sylvia belongs to the alpha matriline and Hannah to the beta. This can only mean that the beta family is attempting to depose the alpha!”

Sixth, knowledge is *independent of sensory modality*: while playback experiments allow us to explore the structure of primates' social knowledge and demonstrate that such knowledge can be acquired through vocalizations alone, social knowledge is doubtless also obtained visually.

These properties of nonhuman primates' social knowledge, while by no means fully human, bear striking resemblances to the meanings we express in language, which are – apparently spontaneously – built up by combining discrete-valued entities in a structured, hierarchical, rule-governed, and open-ended manner (Senghas et al. 2004). This leads to the hypothesis that the internal representations of language meaning in the human brain initially emerged from our pre-linguistic ancestors' knowledge of social relations (Cheney and Seyfarth 1997b; Worden 1998). Social groups of monkeys are comprised of intricate networks of relationships that are simultaneously competitive and cooperative, and in which alliances, friendships, and rivalries cut across both dominance ranks and families. The demands of social life create selective pressures for just the kind of complex, abstract conceptual abilities that are likely to have preceded the earliest forms of linguistic communication. Indeed, as Worden (1998) argues, “no other candidate meaning structure has such a good fit to language meanings.”

We are not suggesting that all of the syntactic properties found in language are present in primate social knowledge. Syntax, in the sense most commonly used by followers of Chomsky, remains a uniquely human trait. Instead, taking an evolutionary perspective, we suggest that the primate mind evolved in an environment characterized by intense social competition, that such competition created selective pressures favoring structured, hierarchical, rule-governed intelligence, and that such social intelligence shares many formal features with linguistic intelligence. If we regard syntax as a succession of computations, then the social knowledge of primates, including their ability to “decode” the propositional information contained in a string of vocalizations, can be seen as a precursor of the kinds of computations found in modern language.

The hypothesis that social cognition acted as a causal engine driving the evolution of mental state attribution, vocabulary growth, and, ultimately, syntax gains credence in part because the alternatives seem less likely. For example, given what we know about the communication of monkeys and apes, it seems unlikely that flexible articulation, and hence the ability to generate new words, would have preceded the conceptual capacity that would have made such articulation adaptive. This argument is supported by research on word learning in children which indicates that words are built upon pre-existing conceptual representations (Hespos and Spelke 2004; Gleitman and Papafragou 2005).

Similarly, it is hard to imagine how the sorts of word learning and simple syntax that we see in young children would have emerged without some rudimentary ability to attribute attention and intent to others (Tomasello 2003). Just

as the social knowledge of monkeys appears to incorporate some of the hierarchical, rule-governed features of human syntax, so does their communication incorporate some rudimentary elements of young children's intention-reading. When responding to vocal signals, monkeys appear to attend simultaneously to signaler identity, the type of call given, the nature of recent interactions with the signaler, and the target of her attention. We have argued that these inferences may represent a first crucial step toward the recognition of other individuals' intentions and motives. The ability to make such inferences would seem to be particularly adaptive in large social groups such as those found in many species of monkeys, where it would be disadvantageous simply to react reflexively to every vocalization that can be heard.

We began this essay by arguing that theories of language evolution cannot focus exclusively on language – they require a broader perspective. To explain the *evolution* of language, one must start not with its most complex, highly evolved features but instead with the simpler properties of communication and cognition that are shared by human and nonhuman primates, since it was these features that set the stage for the emergence of modern language in all of its complexity. We suggest that these features are: the ability of speakers to produce acoustically distinct, individually recognizable vocalizations in specific contexts, the ability of listeners to assign meaning to such calls, and the ability of individuals to organize meaning in ways that allow them to make deductions about events in the world around them.

The existence of such features in the communication and cognition of contemporary nonhuman primates leads us to conclude that the earliest forms of syntax did not emerge *de novo* but instead built upon pre-existing skills. Before hominids produced syntactic utterances, they assigned meaning to other animals' calls and extracted syntactic, rule-governed, propositional information from them. Linguistic-like perception and cognition thus preceded linguistic production.

In its very earliest stages, syntax was adaptive because even the simplest rule-governed two-word utterances allowed individuals to inform, modify, explain, and elaborate in ways that were not possible with single words. The earliest syntactic utterances, however, probably described relations that their speakers already understood, and had a formal structure that grew out of their speakers' knowledge of social relations.

Acknowledgments

We thank Sudha Arunachalam, Thore Bergman, Shermin deSilva, Julia Fischer, Tecumseh Fitch, Lila Gleitman, Marc Hauser, Ray Jackendoff, David January, Ed Kako, John Trueswell, David White, Klaus Zuberbuhler, and two

anonymous reviewers for comments. Research was supported by the National Science Foundation and National Institutes of Health (MH-62249).

University of Pennsylvania

References

- Anderson, J. R., M. Montant, and D. Schmitt (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object choice task. *Behavioural Processes* 37: 47–55.
- Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language* 20: 395–418.
- Bergman, T., J. Beehner, D. L. Cheney, and R. M. Seyfarth (2003). Hierarchical classification by rank and kinship in baboons. *Science* 302: 1234–1236.
- Bickerton, D. (1990). *Language and Species*. Chicago: Univ. of Chicago Press.
- Bloom, P. (2003). *How Children Learn the Meaning of Words*. Cambridge, MA: MIT Press.
- Bloom, P. and L. Markson (1998). Capacities underlying word learning. *Trends in Cognitive Science* 2: 67–73.
- Boyd, R. and J. B. Silk (2000). *How Humans Evolved, 2nd Edition*. New York: W. W. Norton.
- Calvin, W. H. and D. Bickerton (2000). *Lingua Ex Machina: Reconciling Darwin with the Human Brain*. Cambridge, MA: MIT Press.
- Cheney, D. L. and R. M. Seyfarth (1990). *How Monkeys See the World*. Chicago: Univ. of Chicago Press.
- (1997a). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal Behaviour* 54: 409–418.
- (1997b). Why animals don't have language. *The Tanner Lectures on Human Values* 19: 173–210.
- (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behaviour* 58: 67–75.
- (2004). Social complexity and the information acquired during eavesdropping by primates and other animals. In *Animal Communication Networks*, P. McGregor (ed.), 583–603. Cambridge: Cambridge Univ. Press.
- Cheney, D. L., R. M. Seyfarth and J. B. Silk (1995a). The responses of female baboons to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology* 109: 134–141.
- (1995b). The role of grunts in reconciling opponents and facilitating social interactions among adult female baboons. *Animal Behaviour* 50: 249–257.
- Cheney, D. L., R. M. Seyfarth and R. A. Palombit (1996). The function and mechanisms underlying baboon contact barks. *Animal Behaviour* 52: 507–518.
- Clark, A. P. and R. W. Wrangham (1994). Chimpanzee arrival pant-hoots: Do they signify food or status? *International Journal of Primatology* 15: 185–205.
- Crockford, C. and C. Boesch (2003). Context-specific calls in wild chimpanzees, *Pan troglodytes verus*: Analysis of barks. *Animal Behaviour* 66: 115–125.
- Crockford, C., I. Herbinger, L. Vigilant and C. Boesch (2004). Wild chimpanzees have group-specific calls: A case for vocal learning? *Ethology* 110: 221–243.
- Darwin, D. (1871). *The Descent of Man, and Selection in Relation to Sex*. Princeton: Princeton Univ. Press.
- Deacon, T. (1997). *The Symbolic Species*. New York: W. W. Norton.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT/Bradford Books.
- Elowson, M. and C. T. Snowdon (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour* 47: 1267–1277.

- Engh, A. L., R. R. Hoffmeier, D. L. Cheney and R. M. Seyfarth (in press). Who me? Can baboons infer the target of vocalizations? *Animal Behaviour*.
- Fischer J., D. L. Cheney and R. M. Seyfarth (2000) Development of infant baboon responses to female graded variants of barks. *Proceedings of the Royal Society, London, Series B* 267: 2317–2321.
- Fischer, J., K. Hammerschmidt, R. M. Seyfarth and D. L. Cheney (2001a). Acoustic features of female chacma baboon barks. *Ethology* 107: 33–54.
- Fischer, J., M. Metz, D. L. Cheney and R. M. Seyfarth (2001b). Baboon responses to graded bark variants. *Animal Behaviour* 61: 925–931.
- Fischer, J., K. Hammerschmidt, D. L. Cheney and R. M. Seyfarth (2002). Acoustic features of male baboon loud calls: Influence of context, age, and individuality. *Journal of the Acoustical Society of America* 111: 1465–1474.
- Fisher, C. and L. R. Gleitman (2002). Language acquisition. In *Stevens Handbook of Experimental Psychology*, Vol. 3: Learning and Motivation, H. F. Pashler and C. R. Gallistel (eds.), 445–496. New York: Wiley.
- Fitch, W. T. (2004). The evolution of language. In *The Cognitive Neurosciences III*, M. Gazzaniga (ed.), 837–846. Cambridge, MA: MIT Press.
- Fitch, W. T. and M. D. Hauser (1995). Vocal production in nonhuman primates: acoustics, physiology, and functional constraints on “honest” advertisement. *American Journal of Primatology* 37: 191–220.
- Gleitman, L. R. and A. Papafragou, A. (2005). Language and thought. In *Cambridge Handbook of Thinking and Reasoning*, K. Holyoak and R. Morrison (eds.), 633–661. Cambridge: Cambridge University Press.
- Grice, H. (1957). Meaning. *Philosophical Review* 66: 377–388.
- Gyger, M., S. J. Karakashian and P. Marler (1986). Avian alarm-calling: Is there an audience effect? *Animal Behaviour* 34: 1570–1572.
- Hare, B., J. Call, B. Agnetta and M. Tomasello (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59: 771–785.
- Hare, B., J. Call and M. Tomasello (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour* 61: 139–151.
- Hauser, M. D. (1988). How infant vervet monkeys learn to recognize starling alarm calls: The role of experience. *Behaviour* 105: 187–201.
- (1989). Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalizations. *Journal of Comparative Psychology* 103: 149–158.
- (1992). Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production? *Journal of the Acoustic Society of America* 91: 2175–2179.
- (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Herman, L. M., S. A. Kuczaj, and M. D. Holder (1993). Responses to anomalous gestural sequences by a language-trained dolphin: Evidence for processing of semantic relations and syntactic information. *Journal of Experimental Psychology: General* 122: 184–194.
- Hespos, S. J. and E. S. Spelke (2004). Conceptual precursors to language. *Nature* 430: 453–456.
- Hurford, J. D. (1998). Introduction: The emergence of syntax. In *Approaches to the Evolution of Language*, J. R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), 299–304. Cambridge: Cambridge University Press.
- (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26: 261–316.
- Jackendoff, R. (1987). *Consciousness and the Computational Mind*. New York: Basic Books.
- (2002). *Foundations of Language*. Oxford: Oxford Univ. Press.
- Kaminski, J., J. Call and J. Fischer (2004). Word learning in a domestic dog: Evidence for ‘fast mapping’. *Science* 304: 1682–1683.

- Kirby, S. (1998). Fitness and the selective adaptation of language. In *Approaches to the Evolution of Language*, J. R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), 359–383. Cambridge: Cambridge Univ. Press.
- Kitchen, D., R. M. Seyfarth, J. Fischer and D. L. Cheney (2003). Loud calls as indicators of dominance in male baboons. *Behavioral Ecology and Sociobiology* 53: 374–384.
- Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, MA: Harvard Univ. Press.
- (1991). *Uniquely Human: The Evolution of Speech, Thought, and Selfless Behavior*. Cambridge, MA: Harvard Univ. Press.
- Locke, J. (1690). *An Essay Concerning Human Understanding*, A. D. Woozley (ed.). Cleveland: Meridian Books, 1964.
- Manser, M., R. M. Seyfarth, R. M. and D. L. Cheney (2001). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Science* 6: 55–57.
- Marler, P. (1977). The structure of animal communication sounds. In *Recognition of Complex Acoustic Signals*, T. H. Bullock (ed.), 17–35. Berlin: Dahlem Konferenzen.
- Marler, P., C. Evans and M. D. Hauser (1992). Animal signals: Motivational, referential, or both? In *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, H. Papoušek, U. Jürgens, M. Papoušek (eds.), 66–86. Cambridge: Cambridge Univ. Press.
- McGrew, W. C. (1994). Tools compared: The material of culture. In *Chimpanzee Cultures*, R. W. Wrangham, W. C. McGrew, F. B.M. de Waal and P. G. Heltne (eds.), 25–40. Cambridge, MA: Harvard Univ. Press.
- Melnick, D. and M. C. Pearl (1987). Cercopithecines in multi-male groups: Genetic diversity and population structure. In *Primate Societies*, B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham and T. Struhsaker (eds.), 121–134. Chicago: Univ. of Chicago Press.
- Mitani, J. (1993). Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour* 45: 735–746.
- (1996). Comparative studies of African ape vocal behavior. In *Great Ape Societies*, W. C. McGrew, L. F. Marchant, and T. Nishida (eds.), 241–254. Cambridge: Cambridge Univ. Press.
- Mitani, J. and K. L. Brandt (1994). Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology* 96: 233–252.
- Mithen, S. J. (1996). *The Prehistory of the Mind: A Search for the Origins of Art, Religion, and Science*. London: Thames and Hudson.
- Newmeyer, F. (2003). Grammar is grammar and usage is usage. *Language* 79: 682–707.
- Nowak, M. A. and N. L. Komarova (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences* 5: 288–295.
- Nowak, M. A., J. B. Plotkin and V. A.A. Jansen (2000). The evolution of syntactic communication. *Nature* 404: 495–498.
- Owren, M. J. and D. Rendall (1997). An affect-conditioning model of nonhuman primate vocal signaling. In *Perspectives in Ethology, Vol. 12*, M. D. Beecher, D. H. Owings and N. S. Thompson (eds.), 299–346. New York: Plenum Press.
- Owren, M. J., J. A. Dieter, R. M. Seyfarth and D. L. Cheney (1992). “Food” calls produced by adult female rhesus and Japanese macaques, their normally raised offspring, and offspring cross-fostered between species. *Behaviour* 120: 218–231.
- (1993). Vocalizations of rhesus and Japanese macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology* 26: 389–406.
- Owren, M. J., R. M. Seyfarth and D. L. Cheney (1997). The acoustic features of vowel-like grunt calls in chacma baboons: Implications for production processes and function. *Journal of the Acoustical Society of America* 101: 2951–2963.
- Pearce, A. (1985). A review of attempts to condition operantly alloprimate vocalizations. *Primates* 26: 202–213.
- Pepperberg, I. M. (1992). Proficient performance of a conjunctive, recursive task by an African gray parrot (*Psittacus erithacus*). *Journal of Comparative Psychology* 106: 295–305.

- Pinker, S. (1994). *The Language Instinct*. New York: William Morrow and Sons.
- Pinker, S. and P. Bloom (1990). Natural language and natural selection. *Behavioral and Brain Sciences* 13: 713–783.
- Pinker, S. and R. Jackendoff (2005). The faculty of language: What's special about it? *Cognition* 95: 201–236.
- Povinelli, D. and T. J. Eddy (1996). What chimpanzees know about seeing. *Monographs of the Society for Research in Child Development* 61: 1–152.
- Premack, D. (1976). *Intelligence in Ape and Man*. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Premack, D. and G. Woodruff (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 4: 515–526.
- Rendall, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *Journal of the Acoustical Society of America* 113: 3390–3402.
- Rendall, D., P. S. Rodman and R. E. Emond (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour* 51: 1007–1015.
- Rendall, D., R. M. Seyfarth, D. L. Cheney and M. J. Owren (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour* 57: 583–592.
- Rendall, D., D. L. Cheney and R. M. Seyfarth (2000). Proximate factors mediating 'contact' calls in adult female baboons and their infants. *Journal of Comparative Psychology* 114: 36–46.
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist* 43: 151–160.
- Robinson, J. G. (1984). Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus nigrivittatus*. *Behaviour* 90: 46–79.
- Savage-Rumbaugh, E. S. (1986). *Ape Language: From Conditioned Response to Symbol*. New York: Columbia University Press.
- Schusterman, R. J. and K. Krieger (1986). Artificial language comprehension and size transposition by a California sea lion (*Zalophus californianus*). *Journal of Comparative Psychology* 100: 348–355.
- Schusterman, R. J., R., R. Gisiner, B. K. Grimm and E. B. Hanggi (1993). Behavior control by exclusion and attempts at establishing semanticity in marine mammals using match-to-sample paradigms. In *Language and Communication: Comparative Perspectives*, H. Roitblat, L. M. Herman and P. E. Nachtigall (eds), 249–274. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Schusterman, R. J. and D. A. Kastak (1998). Functional equivalence in a California sea lion: relevance to animal social and communicative interactions. *Animal Behaviour* 55: 1087–1095.
- Senghas, A., S. Kita and A. Ozyurek (2004). Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305: 1779–1782.
- Seyfarth, R. M. and D. L. Cheney (1986). Vocal development in vervet monkeys. *Animal Behaviour* 34: 1640–1658.
- (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour* 40: 754–764.
- (1997). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning and Behavior* 25: 249–267.
- (2003a). The structure of social knowledge in monkeys. In *Animal Social Complexity: Intelligence, Culture, and Individualized Societies*, F. de Waal and P. Tyack (eds.), 207–229. Cambridge, MA: Harvard Univ. Press.
- (2003b). Signalers and receivers in animal communication. *Annual Review of Psychology* 54: 145–173.
- 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: 801–803.
- Silk, J. B. (2002). Grunts, girneys, and good intentions: The origins of strategic commitment in nonhuman primates. In *Commitment: Evolutionary Perspectives*, R. Nesse (ed.), 138–157. New York: Russell Sage Press.

- Silk, J. B., D. L. Cheney, and R. M. Seyfarth (1996). The form and function of post-conflict interactions among baboons, *Papio cynocephalus ursinus*. *Animal Behaviour* 52: 259–268.
- Silk, J. B., R. M. Seyfarth and D. L. Cheney (1999). The structure of social relationships among female savanna baboons. *Behaviour* 136: 679–703.
- Snowdon, C. T. (1987). A comparative approach to vocal communication. *Nebraska Symposium on Motivation* 35: 145–200.
- Snowdon, C. T. (1990). Language capacities of nonhuman primates. *Yearbook of Physical Anthropology* 33: 215–243.
- Thompson, R. K.R. (1995). Natural and relational concepts in animals. In *Comparative Approaches to Cognitive Science*, H. Roitblat and J. A. Meyer (eds.), 175–224. Cambridge, MA: MIT Press.
- Tomasello, M. (2003). *Constructing a language: A Usage-based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. and J. Call (1997). *Primate Cognition*. Oxford: Oxford Univ. Press.
- Worden, R. (1998). The evolution of language from social intelligence. In *Approaches to the Evolution of Language*, J. R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), 148–168. Cambridge: Cambridge Univ. Press.
- Zentall, T. R. 1996. An analysis of imitative learning in animals. In *Social Learning in Animals: The Roots of Culture*, C. M. Heyes and B. G. Galef (eds.), 221–244. New York: Academic Press.
- Zuberbuhler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour* 63: 293–299.
- (2003). Referential signaling in non-human primates: Cognitive precursors and limitations for the evolution of language. *Advances in the Study of Behavior* 33: 265–307.
- Zuberbuhler, K., R. Noe and R. M. Seyfarth (1997). Diana monkey long-distance calls: Messages for conspecifics and predators. *Animal Behaviour* 53: 589–604.
- Zuberbuhler, K., D. L. Cheney and R. M. Seyfarth (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology* 113: 33–42.