



Production, usage, and comprehension in animal vocalizations

Robert M. Seyfarth*, Dorothy L. Cheney

Departments of Psychology and Biology, University of Pennsylvania, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:

Accepted 12 October 2009

Available online 26 November 2009

Keywords:

Vocal communication

Call production

Call usage

Response to calls (comprehension)

Mammals

Birds

Nonhuman primates

ABSTRACT

In this review, we place equal emphasis on production, usage, and comprehension because these components of communication may exhibit different developmental trajectories and be affected by different neural mechanisms. In the animal kingdom generally, learned, flexible vocal production is rare, appearing in only a few orders of birds and few species of mammals. Compared with humans, the majority of species produce a limited repertoire of calls that show little modification during development. Call usage is also highly constrained. Unlike humans, most animals use specific call types only in a limited range of contexts. In marked contrast to production and usage, animals' comprehension of vocalizations, as measured by their responses, are highly flexible, modifiable as a result of experience, and show the most parallels with human language. The differences among vocal production, usage, and comprehension create an oddly asymmetric system of communication in which a small repertoire of relatively fixed calls, each linked to a particular context, can nonetheless give rise to an open-ended, highly modifiable, and cognitively rich set of meanings. Recent studies of baboons and eavesdropping songbirds provide two examples.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

In 1871, Charles Darwin drew attention to a dichotomy in the vocal communication of animals that had perplexed philosophers and naturalists for at least 1000 years. In marked contrast to human language, he wrote, animal vocalizations appeared to be involuntary expressions of emotion and movement: "When the sensorium is strongly excited, the muscles of the body are generally thrown into violent action; and as a consequence, loud sounds are uttered, . . . although the sounds may be of no use" (Darwin, 1871/1981: 83). Two pages later, however, Darwin wrote: "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. . . . 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" (1871/1981: 85).

For Darwin, a comparison between human language and the vocal communication of animals necessarily involved the simultaneous study of production (defined as correct articulation, or pronunciation), usage (using vocalizations in the appropriate context), and comprehension (correct understanding of another's vocalizations). This three-pronged approach was essential because, whereas humans were very different from other animals in the do-

main of production, human and nonhuman animals seemed a bit more alike in usage and even more alike in comprehension.

In this paper we follow Darwin's lead and consider production, usage, and comprehension in the vocal communication of animals. Our primary focus is on mammals, because songbirds are the subject of many other papers in this volume. Within the mammals we focus on nonhuman primates, where the most information is available. We also include comparisons with nonprimate mammals and birds.

Modern research on vocal communication in primates supports Darwin's view. In marked contrast to the flexible phonation found in humans, vocal production in monkeys and apes is highly constrained. Each species has a relatively small repertoire of calls whose acoustic features are largely fixed at birth and show little modification during development (see Hammerschmidt and Fischer (2008) for a review). In contrast, as we review below, nonhuman primate vocal usage is somewhat more flexible, and thus more like that found in humans, while comprehension is the most flexible and human-like of all. These generalizations, moreover, apply to many birds and mammals. Throughout the animal kingdom, the ability to modify vocal production depending upon experience is comparatively rare. By contrast, some animals can modify the context in which calls are used, and virtually all species can learn to respond to novel auditory stimuli and change their responses to calls as a result of experience.

For those interested in comparing animal vocal communication with human language, three conclusions emerge. First, flexible vocal production separates humans not only from nonhuman

* Corresponding author. Fax: +1 215 898 7301.

E-mail address: seyfarth@psych.upenn.edu (R.M. Seyfarth).

primates but also from most other mammals and birds. The rarity of learned, modifiable call production throughout most of the Class *Mammalia* suggests that important changes in the mechanisms governing human phonation occurred relatively recently in human evolution, after the divergence of our hominid ancestors from the common ancestors of humans and chimpanzees. Second, whereas an exclusive focus on vocal production reveals clear differences between humans and most other species, a broader examination of production, usage, and comprehension paints a more complicated picture, with both similarities and differences between the two groups. Third, analysis of production, usage, and comprehension draws attention to the very different mechanisms that underlie the behavior of signalers and recipients, even when they are involved in the same communicative event. Many species of birds and mammals have only a small repertoire of acoustically fixed vocalizations. However, because calls are individually distinctive and each call type is predictably linked to a particular social context, this limited call repertoire can nonetheless provide listeners with an open-ended, highly modifiable, and cognitively rich set of meanings, because individuals can continue to learn new sound-meaning pairs throughout their adult lives. As a result, listeners can potentially acquire a huge number of messages from a finite number of call types.

2. The contrast between production, usage, and comprehension

2.1. Production

Monkeys and apes have a small repertoire of calls that show relatively little modification in their acoustic properties during development. The development of vocal production is largely unaffected by variations in auditory experience or rearing (see Hammerschmidt and Fischer (2008) for a recent review).

Although these two features of call production have led many (e.g. Bickerton, 1990) to conclude that call production is also involuntary, this is not accurate. In both the field and the laboratory, nonhuman primates can control whether they produce a vocalization or remain silent. After threatening a subordinate, for example, a female baboon (*Papio hamadryas ursinus*) may or may not give a 'reconciliatory' grunt to her opponent (Cheney & Seyfarth, 1997; Cheney, Seyfarth, & Silk, 1995). When capuchin monkeys (*Cebus capucinus*) find food, they may call or remain silent (Gros-Louis, 2004). Even in highly emotional circumstances like encounters with predators, some individuals give alarm calls at high rates, others call less often, and still others remain silent (Cheney & Seyfarth, 1990). In more controlled laboratory settings, the timing, duration, and rate of calling by monkeys can be brought under operant control (Egnor, Wickelgren, & Hauser, 2007; Pierce, 1985). Clearly, then, primates can control whether they vocalize or not depending upon variations in both the ecological, social, and acoustic environments.

Within a given context, nonhuman primates can also make subtle modifications in the acoustic structure of their calls (reviewed by Hammerschmidt and Fischer (2008), Seyfarth and Cheney (2008)). To cite just one example, wild chimpanzees (*Pan troglodytes*) in Uganda give long, elaborate pant-hoots either alone or in choruses with others. When two individuals have called together several times, the acoustic features of their pant-hoots begin to converge (Mitani & Brandt, 1994; Mitani & Gros-Louis, 1998). Apparently, individuals modify the acoustic structure of their calls depending upon auditory feedback (see also Crockford, Herbinger, Vigilant, & Boesch, 2004).

Many other studies suggest that, whereas the basic acoustic structure of nonhuman primate calls is relatively fixed, the fine spectrotemporal features of vocalizations can be modified (for re-

views see Egnor and Hauser (2004), Hammerschmidt and Fischer (2008), Seyfarth and Cheney (1997a, 2008)). Listeners discriminate among these modifications (e.g. Herbinger, Papworth, Boesch, & Zuberbuhler, 2009; Slocombe, Townsend, & Zuberbuhler, 2008). Some features are more easily modified than others. As Janik and Slater (1997) first pointed out, temporal features like call duration and amplitude are more easily modified than frequency parameters. The distinction between relatively innate and more modifiable components of phonation is important, because it has significant implications for future research on the neurobiology of primate vocal production (see Egnor & Hauser, 2004; Hammerschmidt & Fischer, 2008 for further discussion).

In their relatively fixed vocal production, nonhuman primates are typical of most mammals and even the great majority of birds. In their 1997 review, Janik and Slater found evidence for learned, modifiable vocal production in only three orders of birds, cetaceans, harbor seals, and humans. Although we may yet be surprised by novel evidence of vocal imitation (e.g. Poole, Tyack, Stoeger-Horwarth, & Watwood, 2005) or creative call combinations (Arnold & Zuberbuhler, 2006; Crockford & Boesch, 2003; Zuberbuhler, 2002) in other species, for the moment it appears that the ability to modify vocal production depending upon experience is comparatively rare.

2.2. Usage

Nonhuman primates use acoustically different vocalizations in different social contexts. Some calls, for example, are given only by higher-ranking to lower-ranking individuals, or only when an "audience" is present. Others are given only when approaching an infant or when the group is moving from one area to another (e.g. Gargett, 1971). Many species give acoustically different alarm calls to different classes of predator. Infants and juveniles must therefore acquire correct vocal usage: how to use each call type in the appropriate context. Studies suggest that the mechanisms underlying call usage have a strong genetic component – although not as strong as the mechanisms underlying production.

For example, both rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques produce *coo* and *gruff* calls, but individuals in the two species use them in different ways. Although Japanese macaques can produce *gruffs*, they do so only rarely. When they vocalize during play, they give a *coo*. Conversely, while rhesus macaques can produce *coos*, they give *gruffs* when they play. In a 3-year cross-fostering experiment, two infant rhesus and two infant Japanese macaques were raised in a group of the other species. Each infant was adopted by a mother of the other species and became fully integrated into its new group. Despite their complete social integration and an ability to produce their adoptive species' calls, however, the cross-fostered juveniles continued to use their own species' vocalizations. In the rhesus macaque groups, cross-fostered Japanese macaques gave *coos* during play even though their playmates gave *gruffs*; in the Japanese macaque groups the opposite occurred. The cross-fostered animals behaved as if the link between call and context was difficult to modify (Owren, Dieter, Seyfarth, & Cheney, 1993).

Slightly more flexibility is apparent in the development of vervet monkey (*Chlorocebus aethiops*) alarm calls. Vervets give acoustically different alarm calls to leopards, eagles, and snakes. Each call type elicits a different, adaptive response. Individuals on the ground run into trees when they hear a leopard alarm, look up in the air when they hear an eagle alarm, and peer into the grass around them when they hear a snake alarm (Seyfarth, Cheney, & Marler, 1980). Confronted with a wide variety of potential predators, adult vervets are highly selective, giving 'leopard alarms' to mammalian carnivores, 'eagle alarms' primarily to martial (*Polemaetus bellicosus*) and crowned (*Stephanoetus coronatus*)

eagles, and 'snake alarms' primarily only to pythons (*Python sebae*). Infants and juveniles, by contrast, make many more mistakes, by giving alarm calls to species like warthogs or pigeons that pose no danger to them. Their mistakes, however, are not entirely random. They give leopard alarms almost exclusively to terrestrial mammals, 'eagle alarms' to birds, and 'snake alarms' to reptiles and snake-like objects (Seyfarth & Cheney, 1986; Seyfarth et al., 1980). Vervet infants, in other words, behave as if they are predisposed from birth to divide other species into broadly different classes: predator versus non-predator and, within the former class, terrestrial carnivores, eagles, and snakes. With time and experience they sharpen the relation between each alarm call type and the stimulus that elicits it (Cheney & Seyfarth, 1990). The development of alarm call usage in vervets bears a striking resemblance to the "over-extensions" found in young children's early use of words (e.g. Clark, 2003).

This mixture of relatively innate mechanisms and experience in the development of vocal usage is not unique to vervet alarm calls. It has also been documented in the development of vervet grunts (Seyfarth & Cheney, 1986), intergroup *wrrs* (Hauser, 1989), rhesus macaque *coos* and *gruffs* (Owren et al., 1993), the screams of pigtail macaques (*M. nemestrina*) (Gouzoules & Gouzoules, 1989a, 1989b), the alarm calls of sifakas (*Propithecus verreauxi*, Fichtel & van Schaik, 2006), and the calls of many other primate species (Seyfarth & Cheney, 1997a).

Such results, moreover, are not surprising: there are many reasons why we should not expect vocal usage to be fixed and adult-like at first emergence. To begin with, environments change, and animals that use calls to signal about features of their environment must be able to adjust accordingly. For example, whereas vervets throughout Africa face avian predators, the exact predator species and its physical appearance can vary considerably from one local area to another (e.g. Gargett, 1971). Under these conditions, vervets clearly benefit by using an eagle alarm call that, while closely tied to a particular class of predators, can nonetheless be modified to include new and physically different exemplars within each class. Many other call types are given only to particular individuals, or only to specific individuals in particular circumstances. Call production is also affected by the presence of an audience; for example, female vervets are more likely to give alarm calls when in the presence of their own offspring than when in the presence of an unrelated juvenile (Cheney & Seyfarth, 1990). In these cases it would obviously be impossible for selection to favor completely innate vocal usage, since juveniles require time and experience before they develop close relationships and learn, for example, the individual identities and dominance ranks of others.

2.2.1. Comparison with other mammals

Although data on vocal usage in nonprimate mammals are rare, the few studies that have been conducted suggest that these species are similar to nonhuman primates. In many mammalian species, individuals are generally alert to potential predators but require experience before they can give the appropriate alarm call to different predators (Griffin, Evans, & Blumstein, 2001). The development of vocal usage in the predator-specific alarm calls of young meerkats (*Suricata suricatta*) closely resembles that of vervet monkeys (Hollén, Clutton-Brock, and Manser (2008). And like vervet monkeys (Cheney & Seyfarth, 1990), Belding's ground squirrels (*Spermophilus beldingi*) exhibit an audience effect by giving more calls to predators when close kin are present than when they are not (Sherman, 1977).

2.2.2. Comparison with birds

Compared with the wealth of studies on song production, comparatively little is known about the development of vocal usage in birds. However, the few studies available suggest that birds, like

nonhuman primates and other mammals, exhibit a combination of innate predispositions and more open-ended learning in their use of context-specific vocalizations.

Like nonhuman primates, many birds use acoustically different vocalizations – both different calls and different song types – in different circumstances. Experiments suggest that context-specific vocal usage can be acquired through experience. In a classic study, naïve young blackbirds (*Turdus merula*) learned to mob a harmless stimulus after they had observed another blackbird mobbing it (Curio, Ernst, & Vieth, 1978). We do not know whether blackbirds need experience before they can produce their species-specific mobbing call, but it is clear that their usage of this call is flexible and modifiable. Both blue-winged (*Vermivora pinus*) and yellow (*Dendroica petechia*) warblers have two song types that they use at different times of day and sing at different rates. When young birds are tutored with the two songs types in reversed context and rate of presentation, they duplicate this reversed pattern in their own singing (Kroodsma, 1988; Spector, McKim, & Kroodsma, 1989).

Perhaps the most striking examples of learned, context-specific vocal usage in birds come from studies of song matching. This occurs when a bird modifies its song so that its acoustic features match those of either a local dialect or a neighbor's song (e.g. Beecher, Stoddard, Campbell, & Horning, 1996). Song matching is adaptive because it allows individuals to compete more effectively with their neighbors (Burt, Campbell, & Beecher, 2001; Molles & Vehrencamp, 2001; Vehrencamp, 2001). Young song sparrows (*Melospiza melodia*) need only limited experience to learn their neighbors' song types (Peters, Marler, & Nowicki, 1992). They may learn by eavesdropping or through direct experience (Beecher, Burt, O'Loughlen, Templeton, & Campbell, 2007) and are more likely to learn songs sung by several neighbors than songs sung by a single neighbor (Nordby, Campbell, Burt, & Beecher, 2000).

Nelson and Marler (1994) propose that song matching is achieved through a process of selective attrition and occurs during a period when new song types can no longer be learned. They note that, during the plastic stage of song development, young male song sparrows produce many more song types than are required for a mature repertoire. Over time, some song types are lost, leaving mature birds with a smaller repertoire that remains constant for the rest of the bird's life. Nelson and Marler (1994) showed through experimentation that, during the plastic phase of song learning, young white-crowned sparrows copied (and later crystallized) songs heard from a neighbor that matched one of their own song types, but failed to copy or crystallize songs from a neighbor that did not match one of theirs. Nordby, Campbell, and Beecher (2007) found that during plastic song young song sparrows developed a repertoire of, on average, 10.2 song types, then selectively dropped song types that did not match those of their neighbors. This left them with a crystallized repertoire of, on average, 9.3 song types. Song types that were not shared with one or more of their neighbors were significantly more likely to be dropped than were shared song types.

As with nonhuman primates and other mammals, moderately flexible vocal usage in birds is not surprising. Young birds must learn to recognize their predators, mates, and rivals, all of which are likely to change during an individual's lifetime (Marler, 1977, 1980). Regardless of whether call production is innate or depends on experience, natural selection cannot favor completely innate vocal usage – too much flexibility is required for survival and reproduction. In birds as in nonhuman primates, many different mechanisms may be involved in the developmental transitions that ultimately result in context-specific vocal usage. These include observational learning, eavesdropping, and overproduction coupled with selective attrition, the latter reminiscent of the 'overgeneralization' found in children's early production of words (see above).

2.3. Comprehension

Whenever there is a predictable relation between a particular call type and a specific social context, a vocalization has the potential to inform nearby listeners about objects or events. The underlying mechanisms are irrelevant. A tone that informs a rat about the imminence of a shock, an alarm call that informs a vervet about the presence of an eagle, or a scream that informs a baboon that her offspring is involved in a fight all have the potential to provide a listener with precise information if they are predictably associated (Rescorla, 1988) with a narrow range of events. The widely different mechanisms that lead to this association have no effect on the signal's potential to inform (Seyfarth & Cheney, 2003).

But while they have this potential, do animal vocalizations really provide listeners with 'information'? The study of 'comprehension' in animal communication is fraught with both practical and conceptual difficulties because we cannot interview our subjects. As a result, the only way to determine the 'meaning' of a signal is to examine the recipient's response – a very crude measure of whether the recipient has or has not acquired information.

In some cases, animals respond in qualitatively different ways to different vocalizations, and the characteristics of their responses suggest that each call type conveys specific information. Consider, for example, the responses of vervet monkeys (described above), Diana monkeys (*Cercopithecus diana*), and meerkats to playbacks of different alarm calls. When Diana monkeys hear the growl of a leopard, they respond with their own 'leopard alarm'; when they hear the shriek of an eagle, they give their own 'eagle alarm' (Zuberbuhler, Cheney, & Seyfarth, 1999). When meerkats hear a terrestrial predator alarm they approach the nearest burrow; an eagle alarm causes them to crouch and scan the sky; and a 'recruitment call' (given to snakes) causes them to approach the caller, tail raised, and investigate the area (Manser, Bell, & Fletcher, 2001).

In many other cases, the lack of qualitatively different responses complicates the interpretation of call meaning. Two methods help to circumvent this problem, at least partially. First, if observations suggest that some generic response, like looking toward the speaker, will be the only reaction elicited by a playback experiment, scientists can design a matched pair of trials, alike in all but one respect, and then compare subjects' responses under two different conditions. A consistent difference in the duration of response may permit inferences about the different sorts of information conveyed by different calls, or by the same call under different conditions. Alternatively, subjects may show no immediate response to playback of a call, but their subsequent behavior may nonetheless be affected. Having heard a particular call from individual X, for example, a subject may be more likely to approach X in the next 30 min than if no call, or the call of a different individual, had been played. Such longer-term changes in behavior also allow one to make inferences, albeit indirect, about the meaning of a specific vocalization (Seyfarth and Cheney (1997b), for examples see Cheney and Seyfarth (2007)).

Despite the indirect way in which it is measured, we can nonetheless state with confidence that nonhuman primates' comprehension of vocalizations is very different from their production and usage. Whereas production and usage are highly constrained, comprehension is more flexible, open-ended, and can be modified by experience.

Three month-old vervet monkeys, for example, respond to playback of leopard, eagle, and snake alarm calls by running to their mothers or showing some other, often inappropriate, reaction. They require several months' more experience before they respond to the different alarm calls in an appropriate, adult-like manner (Seyfarth & Cheney, 1986; Seyfarth et al., 1980). Similar gradual development has been documented in the responses of Goeldi's monkeys (*Callimico goeldi*), Japanese macaques, bonnet macaques

(*M. radiata*), baboons, and sifakas to their predator-specific alarm calls (Fichtel & van Schaik, 2006; Fischer, Cheney, & Seyfarth, 2000; Masataka, 1983a, 1983b; Ramachandran & Coss, 2000), and in the emergence of individual recognition among Japanese and Barbary macaques (*M. sylvanus*) (Fischer, 2004; Masataka, 1985).

In their natural habitats, vervets, baboons, and other primates learn to recognize the alarm calls of other species of primates, birds, and ungulates, even though these calls are acoustically very different from their own (Fichtel, 2004; Hauser, 1988; Seyfarth & Cheney, 1990; Zuberbuhler, 2001). They also learn to recognize the calls of their predators (Hauser & Wrangham, 1990; Zuberbuhler et al., 1999). And throughout their lives monkeys and apes must continually learn to identify the voices of individuals who join their group or are born into it. Primates in laboratories readily learn to recognize the voices of their different caretakers and to associate different sounds, like the rattling of keys or the beep of a card-swipe, with impending events that may be good (the delivery of food) or bad (the visit of a veterinarian). In the cross-fostering experiments described earlier, cross-fostered subjects learned to recognize their foster mothers' calls – and the foster mothers learned to recognize theirs – even in contexts in which the two species used acoustically different vocalizations (Seyfarth & Cheney, 1997a).

2.3.1. Comparison with other mammals

Flexible, open-ended comprehension of vocalizations is found in many other mammals. For example, when young Belding's ground squirrels first emerge from their burrows, they do not seem to discriminate among the acoustically different alarm calls that adults give to different predators. However, discrimination emerges rapidly over the next 4 weeks (Mateo, 1996a, 1996b), and may be facilitated by contact with adults (Mateo & Holmes, 1997, 1999; for other examples see Hanson & Coss, 2001 for California ground squirrels (*Spermophilus beecheyi*), Sloan and Hare (2006) for Richardson's ground squirrels (*Spermophilus richardsonii*); Hollén & Manser, 2006 for suricates). As in monkeys and apes, both the recognition of other species' alarm calls (e.g. Blumstein & Armitage, 1997; Randler, 2006; Russ, Jones, Mackie, & Racey, 2004; Schmidt, Lee, Ostfield, & Sieving, 2008), the recognition of group-specific calls (bats, *Phyllostomus hastatus*: Boughman, 1997) and the recognition and selective response to particular individuals (e.g. dolphins (*Tursiops truncatus*): Watwood, Owen, Tyack, and Wells (2005), see Snowdon (1986) for review) are widespread among mammals. Experience is almost certainly required for the development of these recognition systems.

Perhaps the most striking examples of open-ended call comprehension combined with restricted call production come from studies in which individual animals have been raised among humans and learned to respond to human speech. Consider, for example, Rico, a border collie who learned the names of more than 200 different toys (Kaminski, Call, & Fischer, 2004). Rico was able to learn and remember the names of new toys by process of exclusion, or "fast mapping", and – like small children – used gaze and attention to guide word learning. But of course Rico never learned to say any of the words he learned. In this respect, his limited vocal production and extensive comprehension are similar to those of human-trained sea lions (*Zalophus californianus*) (Schusterman, Reichmuth Kastak, & Kastak, 2002), and dolphins (Herman, Pack, & Morrel-Samuels, 1993). The only studies of human-raised animals in which production has equaled comprehension are those that have taken advantage of an animal's natural ability to mimic speech (Pepperberg, 2002), or trained individuals to use an artificial system of sign production that involved mechanisms of sign production that were less constrained than the ones they would normally develop (e.g. signing chimpanzees, Savage-Rumbaugh, 1986; Terrace, 1979).

2.3.2. Comparison with birds

In some cases, birds' responses to auditory stimuli have a strong innate component. From a very young age, male songbirds attend preferentially to their own species' song as opposed to the song of another species (Dooling & Searcy, 1980; King & West, 1977; Nelson & Marler, 1993). Both male and female zebra finches (*Taeniopygia guttata*) reared in auditory isolation show a preference for conspecific song (Braaten & Reynolds, 1999). Yet even these early, innate preferences are not entirely fixed. Although male birds are typically biased toward learning their own species' song, they will also learn to sing all or part of another species' song if given the opportunity (e.g. Clayton, 1987; Eales, 1987; Marler & Peters, 1981). And while female birds may begin life with a general preference for their own species' song or a specific local dialect, exposure to a particular song type or interaction with a specific individual singer can create stable, repeatable preferences that last for months or even years (Freeberg, 2000; Riebel, 2000; Slater, Eales, & Clayton, 1988; West, King, White, Gros-Louis, & Freed-Brown, 2006). Sewall and Hahn (2009) studied the use of contact calls in different ecomorphs of the crossbill (*Loxia curvirostra*). Individuals in each ecomorph produce a contact calls whose acoustic features are subtly different from all others. Adult crossbills typically respond only to calls of their own acoustic subtype. However, experimental pairing with a bird from a different subtype erodes this preference, causing birds to respond with equal strength to all of the different call subtypes (for similar data on matched calling see the references in Sewall and Hahn; for similar experimental results with starlings (*Sturnus vulgaris*) see Gentner & Hulse, 2000).

Conditioning experiments have shown that birds can learn to discriminate among a wide variety of auditory stimuli and respond differently to them (e.g. Adret, 1993 and references therein). In birds like parrots and mynahs, which mimic many sounds, the ability to form an association between a particular auditory stimulus and a specific response seems almost completely unconstrained. In other species, the breadth and flexibility of responses to auditory stimuli is unknown. Data on song matching, however, suggest that within this particular domain of interaction many songbirds have at least as much flexibility as nonhuman primates. Territorial males must learn to recognize the song types sung by their different neighbors and to respond with a matching song type of their own, much as a Diana monkey must learn to recognize the growl of a leopard and to respond with its own 'leopard alarm' call (Zuberbuhler et al., 1999).

3. Constrained production, open-ended comprehension

When a species' calls are individually distinctive and different call types are predictably associated with specific social contexts, listeners confront an environment in which there are statistical regularities linking caller, call type, social context, and behavior. Listeners are sensitive to these associations. From them, they extract whatever information is relevant, using the information acquired "to represent their environment" (Rescorla, 1988) – that is, to form some kind of 'mental picture' of events taking place around them, even if they cannot see precisely what is happening. The result is a communication system in which a constrained, finite number of calls gives rise to an almost limitless number of messages. Here are two examples.

3.1. Case study #1: the information acquired from vocalizations by baboons

Baboons live throughout the savannah woodlands of Africa in groups of 50–150 individuals. Although 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 (Silk, Altmann, & Alberts, 2006a, 2006b; Silk, Seyfarth, & Cheney, 1999). Females can be ranked in a stable, linear dominance hierarchy that determines priority of access to resources, and daughters acquire ranks similar to those of their mothers. Baboon social structure can therefore be described as a hierarchy of matriline, in which all members of one matriline (for example, matriline B) outrank or are outranked by all members of another (for example, matriline C and A, respectively). Ranks within matriline are as stable as those between matriline (for example, A1 > A2 > A3 > B1 > B2 > C1 > C2) (Cheney & Seyfarth, 2007).

Throughout the day, baboons hear other group members giving vocalizations to each other. Calls are individually distinctive (e.g. Owren, Seyfarth, and Cheney (1997), reviewed in Cheney and Seyfarth, (2007)). Some interactions involve aggressive competition; for example, when a higher-ranking animal gives a series of threat-grunts to a lower-ranking animal and the latter screams. Threat-grunts are aggressive vocalizations given by higher-ranking to lower-ranking individuals, whereas screams are submissive signals, given primarily by lower- to higher-ranking individuals. A threat-grunt-scream sequence, therefore, provides information not only about the identities of the opponents involved but also about who is threatening whom. Baboons are very sensitive to both types of information. In playback experiments, listeners respond with apparent surprise to sequences of calls that appear to violate the existing dominance hierarchy. Whereas they show little response upon hearing the sequence "B2 threat-grunts and C3 screams", they respond strongly – by looking toward the source of the call – when they hear "C3 threat-grunts and B2 screams". Apparent between-family rank reversals (C3 threat-grunts and B2 screams) elicit a stronger response than do within-family rank reversals (C3 threat-grunts and C1 screams) (Bergman, Beehner, Cheney, & Seyfarth, 2003).

A baboon who ignores the sequence "B2 threat-grunts and C3 screams" but responds strongly when she hears "C3 threat-grunts and B2 screams" reveals, by her responses, that she recognizes the identities of both participants, their relative ranks, and their family membership. She also acts as if she assumes that the threat-grunt and scream have occurred together not by chance, but because one vocalization caused the other to occur. Without this assumption of causality there would be no violation of expectation when B2's scream and C3's threat-grunt occurred together. Baboons' ability to deduce a social narrative from a sequence of sounds reveals a rich cognitive system in which listeners extract a large number of complex, nuanced messages from a relatively small, finite number of signals. A baboon who understands that "B2 threat-grunts and C3 screams" is different from "C3 threat-grunts and B2 screams" can make the same judgment for all possible pairs of group members as well as any new individuals who may join. Flexible, open-ended comprehension results in a rich system of social cognition in which discrete-valued traits (individual identity, matrilineal kin, rank) are combined in rule-governed ways to create a representation of social structure whose underlying cognitive basis shares some parallels with human language (Cheney & Seyfarth, 2007, chap. 10–11).

3.2. Case study #2: the information acquired from vocalizations by eavesdropping songbirds

Peake, Terry, Dabelsteen, and McGregor (2002) studied the information acquired by great tits (*Parus major*) from listening to song contests between putative neighbors. Male tits defend their territory by singing; challengers sing in reply. When an intruder begins to sing before a male has completed his song, this signals a willingness to escalate: such "overlapping" is correlated with

increased aggression by the overlapper (e.g. [Hultsch & Todt, 1982](#)). When an intruder waits until the resident has completed his song or allows the resident to overlap him, escalated aggression is less likely.

[Peake et al. \(2002\)](#) 'introduced' an 'intruder' (actually, a speaker set up for interactive playback) to a territorial male subject. The intruder, A, either played an aggressive role and overlapped the subject's song or played a less aggressive role and waited for one second after the subject had finished before he began signing. Alternatively, A allowed the subject to overlap him. Next, an interaction was simulated between A and another bird, B (again represented by a loudspeaker), in an area that was adjacent to but outside the subject's territory. Here either A or B played the aggressive role by overlapping the other bird's song. These two simulated interactions created four treatment types that could, in theory, provide the subject with different sorts of information about B's status relative to his own. In two cases the information was ambiguous: either A was aggressive to both B and the subject or A received aggression from both. In two other cases the information was not ambiguous: either A was highly aggressive to the subject but received aggression from B, indicating that B was of high status relative to the subject, or A received aggression from the subject but was highly aggressive to B, indicating that B was of low status relative to the subject.

[Peake et al. \(2002\)](#) then simulated an intrusion of B onto the subject's territory. When compared with their reaction to males that were of relatively low status, subjects showed a threefold increase in singing to high status males or males for whom relative status was ambiguous. Subjects appeared to combine the information acquired from two previous interactions, one in which they took part and one in which they were only eavesdroppers. The combination of information acquired from both sources then determined their response to B's intrusion. Great tits may have a relatively limited repertoire of vocalizations, but, like baboons, their calls are individually distinctive and given according to predictable 'rules'. As a result, a limited, finite repertoire can nonetheless convey an enormous variety of meanings.

4. Discussion

A review of production, usage, and response in the vocal communication of animals yields the following general conclusions:

1. In the animal kingdom generally, learned, flexible vocal production is relatively rare, appearing in only a few orders of birds and few species of mammals. Compared with humans, the great majority of animal species have a limited vocal repertoire of calls that are adult-like in their acoustic properties when they first appear and show little modification during development. This is not to say that call production is completely innate: in most species, individuals can voluntarily produce calls or remain silent and can make some modification in the timing of call production depending upon experience.
2. Because constraints on call production necessarily constrain an animal's ability to use a specific call in a particular context, call usage is also sharply limited. In contrast to humans' very flexible use of different speech sounds in different circumstances, most animals use specific call types only in a limited range of contexts. The link between call type and context is often difficult to sever. Nonetheless, some flexibility in call usage is apparent, leading to some limited parallels between human and nonhuman species. For example, as they develop the ability to give particular alarm call types only to specific predators, young vervet monkeys and young suricates make errors that resemble the over-extensions of word meaning found in human children.
3. Because animals' responses to vocalizations often do not involve call production, this feature of their communication is the least innate, the most modifiable as a result of experience, and shows many more parallels with human language. Many animals are born with little or no predisposition to respond to particular call types in specific ways, but develop the ability to do so with age and experience. In these species there is an almost completely open-ended ability to associate particular sounds with specific individuals, features of the environment, or events. In other species, although young individuals are predisposed to respond selectively to some vocalizations and not others, this early bias can be modified as a result of experience.
4. As Darwin was one of the first to note, the different developmental trajectories of vocal production, usage, and response create an oddly asymmetric system of communication in which a small repertoire of relatively fixed calls, each linked to a particular context, can nonetheless give rise to an open-ended, highly modifiable, and cognitively rich set of meanings. Recent studies of baboons and eavesdropping songbirds provide examples of species in which constrained vocal production and usage by signalers provides listeners with extensive information about their social environment.

The striking difference between production and comprehension is puzzling, because producers are also perceivers: why should an individual who can deduce an almost limitless number of meanings from the calls of others be able to produce only a limited number of calls of her own? The difference may arise because call production depends on mechanisms of phonation, which are largely innate, whereas comprehension depends on mechanisms of learning, including classical and operant conditioning, which are considerably more malleable. But this explanation offers no answer to a crucial question: Why has natural selection so rarely acted to favor flexible vocal production? Here we offer some speculations as they apply to human and nonhuman primates.

At some point in our evolutionary history – probably after the divergence of the evolutionary lines leading to chimpanzees and bonobos on the one hand and humans on the other ([Enard et al., 2002](#)) – our ancestors developed much greater control over the physiology of vocal production. As a result, vocal output became both more flexible and considerably more dependent on auditory experience and imitation ([Fitch, 2007](#); [Lieberman, 1991](#)). What selective pressures might have given rise to these physiological changes?

Vocal communication in nonhuman primates lacks three features that are abundantly present in human language: the ability to generate new words, lexical syntax, and a complete theory of mind. By the latter we mean the ability of both speakers and listeners to make attributions about each others' beliefs, knowledge, and other mental states when communicating with each other ([Grice, 1957](#)). These are the simplest, most basic features that distinguish human and nonhuman primate vocal production, and it is with these traits that speculations about the evolution of language must start. At the earliest stages of language evolution we need not worry about the more complex properties of language that probably came later – properties like case, tense, subject-verb agreement, open- and closed-class items, recursion, long-distance dependency, subordinate clauses, and so on.

How might the ability to generate new words, lexical syntax, and a theory of mind have evolved: simultaneously, in response to the same selective pressures, or more serially, in some particular order? We propose that the evolution of a theory of mind preceded language, creating the selective pressures that gave rise to the ability to generate new words and lexical syntax, and to the flexibility in vocal production that these two traits would have required

(Cheney & Seyfarth, 2005, 2007). We make this argument on both empirical and theoretical grounds.

Empirically, there is no evidence in nonhuman primates for anything close to the large vocal repertoire we find even in very young children. Similarly, nonhuman primates provide few examples of lexical syntax. Recent work by Zuberbuhler and colleagues on the alarm calls of forest monkeys provide intriguing evidence that the presence of one call type can ‘modify’ the meaning of another (Arnold & Zuberbuhler, 2006; Zuberbuhler, 2002), and a study by Crockford and Boesch (2003) suggests that a call combination in chimpanzees may carry new meaning that goes beyond the meaning of the individual calls themselves, but these rare exceptions meet few of the definitions of human syntax.

By contrast, there is growing evidence that both Old World monkeys (Cheney & Seyfarth, 2007; Engh, Hoffmeier, Cheney, & Seyfarth, 2006; Flombaum & Santos, 2005) and apes (Buttelmann, Carpenter, Call, & Tomasello, 2007; Hare, Call, & Tomasello, 2001; Tomasello, Carpenter, Call, Behne, & Moll, 2005) may possess rudimentary abilities to attribute motives or knowledge to others, and engage in simple forms of shared attention and social referencing. Taken together, these data suggest that a rudimentary theory of mind appeared among primates long before flexible vocal production, the ability to generate new words, and lexical syntax.

More theoretically, we suggest that the evolution of a theory of mind acted as a prime mover in the evolution of language because, while it is easy to imagine a scenario in which a rudimentary theory of mind preceded and provided the impetus for the evolution of large vocabularies and syntax, any alternative sequence of events seems less likely.

Consider, for example, the course of word learning in children. Beginning as early as 9–12 months, children exhibit a nascent understanding of other individuals’ motives, beliefs, and desires, and this skill forms the basis of a shared attention system that is essential for early word learning (Bloom & Markson, 1998; Tomasello, 2003). One-year-old children seem to understand that words can be mapped onto objects and actions. Crucially, this understanding is accompanied by a kind of “social referencing” in which the child uses other people’s direction of gaze, gestures, and emotions to assign labels to objects (Baldwin (1991), reviewed in Fisher and Gleitman (2002), Pinker (1994)). Gaze and attention also facilitate word learning in dogs and other animals. Children, however, rapidly surpass the simpler forms of shared attention and word learning demonstrated by animals. Long before they begin to speak in complete sentences, young children develop implicit notions of objects and events, actors, actions, and those that are acted upon. As Fisher and Gleitman (2002) argue, these “conceptual primitives” provide children with a kind of “conceptual vocabulary onto which the basic linguistic elements (words and structures) are mapped.” Moreover, in contrast to monkeys, apes, and other animals, 1-year-old children are highly motivated to share what they know with others (Tomasello & Carpenter, 2007). While animals are concerned with their own goals and knowledge, young children are motivated to make their thoughts and knowledge publically available.

The acquisition of a theory of mind thus creates a cognitive environment that drives the acquisition of new words and new grammatical skills. Indeed, results suggest that children could not increase their vocabularies or learn grammar as rapidly as they do if they did not have some prior notion of other individuals’ mental states (Fisher & Gleitman, 2002; Pinker, 1994; Tomasello, 2003). In sum, data on children provide an excellent illustration of how a theory of mind can drive language development.

By contrast, it is much more difficult to imagine how our ancestors could have learned new words or grammatical rules if they were unable to attribute mental states to others. 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, for example, clearly distinguish between a sequence of calls indicating that Sylvia is threatening Hannah, as opposed to Hannah is threatening Sylvia. Nor does the lack of syntax arise because of an inability to mentally represent descriptive verbs, modifiers, or prepositions. In captivity, a variety of animals, including dolphins (Herman et al., 1993), sea lions (Schusterman & Krieger, 1986), and African gray parrots (Pepperberg, 1992), can be taught to understand and in some cases even to produce verbs, modifiers, and prepositions. Even in their natural behavior, nonhuman primates and other animals certainly seem capable of *thinking* in simple sentences, but the ability to think in sentences does not motivate them to *speak* in sentences. Their knowledge remains largely private.

This limitation may arise because nonhuman primates and other animals cannot distinguish between what they know and others know and cannot recognize, for example, that an ignorant individual might need to have an event explained to them. As a result, although they may mentally tag events as argument structures, they fail to map these tags into a communicative system in any stable or predictable way. Because they cannot attribute mental states like ignorance to others, and are unaware of the causal relation between behavior and beliefs, monkeys and apes do not actively seek to explain or elaborate upon their thoughts. As a result, they are largely incapable of inventing new words or of recognizing when thoughts should be made explicit.

We suggest, then, that long before our ancestors spoke in sentences, they had a language of thought in which they represented the world – and the meaning of call sequences – in terms of actors, actions, and those who are acted upon. The linguistic revolution occurred when our ancestors began to express this tacit knowledge, and to use their cognitive skills in speaking as well as listening. The prime mover behind this revolution was a theory of mind that had evolved to the point where its possessors did not just recognize other individuals’ goals, intentions, and even knowledge – as monkeys and apes already do – but were also motivated to share their own goals, intentions, and knowledge with others. Whatever the selective pressures that prompted this change, it led to a mind that was motivated to make public thoughts and knowledge that had previously remained private. The evolution of a theory of mind spurred the evolution of words and grammar. It also provided the selective pressure for the evolution of the physiological adaptations that enabled vocal modifiability.

References

- Adret, P. (1993). Operant conditioning, song learning, and imprinting to taped song in the Zebra finch. *Animal Behaviour*, 46, 149–159.
- Arnold, K., & Zuberbuhler, K. (2006). Language evolution: Compositional semantics in primate calls. *Nature*, 441, 303.
- Baldwin, D. (1991). Infants’ contribution to the achievement of joint reference. *Child Development*, 92, 875–890.
- Beecher, M. D., Stoddard, P. K., Campbell, S. E., & Horning, C. L. (1996). Repertoire matching between neighbouring song sparrows. *Animal Behaviour*, 51, 917–923.
- Beecher, M. D., Burt, J. M., O’Loughlin, A. L., Templeton, C. N., & Campbell, S. E. (2007). Bird song learning in an eavesdropping context. *Animal Behaviour*, 73, 929–935.
- Bergman, T., Beehner, J., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, 302, 1234–1236.
- Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.
- Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Science*, 2, 67–73.
- Blumstein, D., & Armitage, K. B. (1997). Alarm-calling in yellow-bellied marmots. I. The meaning of situationally specific calls. *Animal Behaviour*, 53, 143–171.
- Boughman, J. W. (1997). Greater spear-nosed bats give group-distinctive calls. *Behavioral Ecology & Sociobiology*, 40, 61–70.
- Braaten, R. F., & Reynolds, K. (1999). Auditory preference for conspecific song in isolate-reared zebra finches. *Animal Behaviour*, 58, 105–111.
- Burt, J. M., Campbell, S. E., & Beecher, M. D. (2001). Song type matching as threat: A test using interactive playback. *Animal Behaviour*, 62, 1163–1170.

- Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated apes imitate rationally. *Developmental Science*, 10, 31–38.
- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world*. Chicago: University of Chicago Press.
- Cheney, D. L., & Seyfarth, R. M. (1997). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal Behaviour*, 54, 409–418.
- Cheney, D. L., & Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *Linguistic Review*, 22, 135–159.
- Cheney, D. L., & Seyfarth, R. M. (2007). *Baboon metaphysics: The evolution of a social mind*. Chicago: University of Chicago Press.
- Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The responses of female baboons to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology*, 109, 134–141.
- Clark, E. (2003). *First language acquisition*. Cambridge: Cambridge University Press.
- Clayton, N. (1987). Song learning in cross-fostered zebra finches: A re-examination of the sensitive phase. *Behaviour*, 102, 67–81.
- Crockford, C., & Boesch, C. (2003). Context-specific calls in wild chimpanzees (*Pan troglodytes verus*): Analysis of barks. *Animal Behaviour*, 66, 115–125.
- Crockford, C., Herbinger, L., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees have group-specific calls: A case for vocal learning? *Ethology*, 110, 221–243.
- Curio, E., Ernst, U., & Vieth, W. (1978). The adaptive significance of avian mobbing. II. Cultural transmission of enemy recognition in blackbirds: Effectiveness and some constraints. *Zeitschrift für Tierpsychologie*, 48, 184–202.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. Princeton: Princeton University Press.
- Dooling, R., & Searcy, M. H. (1980). Early perceptual selectivity in the swamp sparrow. *Developmental Psychobiology*, 13, 499–506.
- Eales, L. A. (1987). Do zebra finch males that have been raised by another species still tend to select a conspecific tutor? *Animal Behaviour*, 35, 1347–1355.
- Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Cognitive Science*, 27, 649–654.
- Egnor, S. E. R., Wickelgren, J., & Hauser, M. D. (2007). Tracking silence. Adjusting vocal production to avoid acoustic interference. *Journal of Comparative Physiology A*, 193, 477–483.
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., et al. (2002). Molecular evolution of FOXP2: A gene involved in speech and language. *Nature*, 418, 869–872.
- Engh, A. E., Hoffmeier, R. R., Cheney, D. L., & Seyfarth, R. M. (2006). Who, me? Can baboons infer the target of vocalizations? *Animal Behaviour*, 71, 381–387.
- Fichtel, C. (2004). Reciprocal recognition of sifaka (*Propithecus verreauxi*) and redfronted lemur (*Eulemur fulvus rufus*) alarm calls. *Animal Cognition*, 7, 45–52.
- Fichtel, C., & van Schaik, C. (2006). Semantic differences in sifaka (*Propithecus verreauxi*) alarm calls: A reflection of genetic or cultural variants? *Ethology*, 112, 839–849.
- Fischer, J. (2004). Emergence of individual recognition in young macaques. *Animal Behaviour*, 67, 655–661.
- Fischer, J., Cheney, D. L., & Seyfarth, R. M. (2000). Development of infant baboon responses to female graded variants of barks. *Proceedings of the Royal Society, London, B*, 267, 2317–2321.
- Fisher, C., & Gleitman, L. R. (2002). Language acquisition. In H. F. Pashler & C. R. Gallistel (Eds.), *Handbook of experimental psychology. Learning and motivation* (Vol. 3, pp. 445–496). New York: Stevens Wiley.
- Fitch, W. T. (2007). The evolution of language: A comparative perspective. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.
- Flombaum, J. L., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15, 447–452.
- Freeberg, T. M. (2000). Culture and courtship in vertebrates: A review of social learning and transmission of courtship systems and mating patterns. *Behavioural Processes*, 51, 177–192.
- Gargett, V. (1971). Some observations on black eagles in the Matopos, Rhodesia. *Ostrich (Supplement)*, 9, 91–124.
- Gentner, T. Q., & Hulse, S. (2000). Female European starling preference and choice for variation in conspecific male song. *Animal Behaviour*, 59, 443–458.
- Gouzoules, H., & Gouzoules, S. (1989a). Design features and developmental modification of pigtail macaque (*Macaca nemestrina*) agonistic screams. *Animal Behaviour*, 37, 383–401.
- Gouzoules, H., & Gouzoules, S. (1989b). Sex differences in the acquisition of communicative competence by pigtail macaques (*Macaca nemestrina*). *American Journal of Primatology*, 19, 163–174.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.
- Griffin, A. S., Evans, C. S., & Blumstein, D. L. (2001). Learning specificity in acquired predator recognition. *Animal Behaviour*, 62, 577–589.
- Gros-Louis, J. (2004). The function of food-associated calls in white-faced capuchin monkeys, *Cebus capucinus*, from the perspective of the signaler. *Animal Behaviour*, 67, 431–440.
- Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In U. Griebel & K. Oller (Eds.), *The evolution of communicative creativity: From fixed signals to contextual flexibility* (pp. 93–119). Cambridge, MA: MIT Press.
- Hanson, M. T., & Coss, Richard G. (2001). Age differences in the response of California ground squirrels (*Spermophilus beecheyi*) to conspecific alarm calls. *Ethology*, 107, 259–275.
- Hare, B., Call, J., & Tomasello, M. (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.
- Hauser, M. D. (1989). Ontogenetic changes in the comprehension of vervet monkey vocalizations. *Journal of Comparative Psychology*, 103, 149–158.
- Hauser, M. D., & Wrangham, R. W. (1990). Recognition of predator and competitor calls in nonhuman primates and birds: A preliminary report. *Ethology*, 86, 116–130.
- Herbinger, I., Papworth, S., Boesch, C., & Zuberbühler, K. (2009). Vocal, gestural, and locomotor responses of wild chimpanzees to familiar and unfamiliar intruders: A playback study. *Animal Behaviour*, doi:10.1016/j.anbehav.2009.09.010.
- Herman, L. M., Pack, A. A., & Morrel-Samuels, P. (1993). Representational and conceptual skills of dolphins. In H. L. Roitblat, L. M. Herman, & P. E. Nachtigall (Eds.), *Comparative cognition and neuroscience* (pp. 403–442). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hollén, L. I., Clutton-Brock, T. H., & Manser, M. B. (2008). Ontogenetic changes in alarm call production and usage in meerkats (*Suricata suricatta*): Adaptations or constraints. *Behavioral Ecology & Sociobiology*, 62, 821–829.
- Hollén, Linda, I., & Manser, M. B. (2006). Ontogeny of alarm call responses in meerkats, *Suricata suricatta*: The roles of age, sex, and nearby conspecifics. *Animal Behaviour*, 72, 1345–1353.
- Hultsch, H., & Todt, D. (1982). Temporal performance roles during vocal interactions in nightingales (*Luscinia megarhynchos*). *Behavioral Ecology & Sociobiology*, 11, 253–260.
- Janik, V. W., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science*, 304, 1682–1683.
- King, A. P., & West, M. J. (1977). Species identification in the North American cowbird: Appropriate responses to abnormal song. *Science*, 195, 1002–1004.
- Kroodtsma, D. (1988). Song types and their use: Developmental flexibility of the male blue-winged warbler. *Ethology*, 79, 235–247.
- Lieberman, P. (1991). *Uniquely human*. Cambridge, MA: Harvard University Press.
- Manser, M. B., Bell, M. B., & Fletcher, L. (2001). The information that receivers extract from alarm calls in suricates. *Proceedings of the Royal Society, London, B*, 268, 2485–2491.
- Marler, P. (1977). Development and learning of recognition systems. In T. H. Bullock (Ed.), *Recognition of complex acoustic signals* (pp. 77–96). Berlin: Dahlem Konferenzen.
- Marler, P. (1980). Comparative perspectives on ethology and behavioral development. In M. Bornstein (Ed.), *Comparative methods in psychology: Ethological, developmental, and cross-cultural perspectives* (pp. 135–151). Hillsdale NJ: Lawrence Erlbaum.
- Marler, P., & Peters, S. (1981). Birdsong and speech: Evidence for special processing. In P. Eimas & J. Miller (Eds.), *Perspectives on the study of speech* (pp. 75–112). Hillsdale NJ: Lawrence Erlbaum.
- Masataka, N. (1983a). Psycholinguistic analyses of alarm calls of Japanese monkeys. *American Journal of Primatology*, 5, 111–125.
- Masataka, N. (1983b). Categorical responses to natural and synthesized alarm calls in Goeldi's monkeys (*Callimico goeldi*). *Primates*, 24, 40–51.
- Masataka, N. (1985). Development of vocal recognition of mothers in infant Japanese monkeys. *Developmental Psychobiology*, 18, 107–114.
- Mateo, J. (1996a). Early auditory experience and the ontogeny of alarm call discrimination in Belding's ground squirrels (*Spermophilus beldingi*). *Journal of Comparative Psychology*, 110, 115–124.
- Mateo, J. (1996b). The development of alarm call response behavior in free-living juvenile Belding's ground squirrels. *Animal Behaviour*, 52, 489–505.
- Mateo, J., & Holmes, W. (1997). Development of alarm call response behavior in juvenile Belding's ground squirrels: The role of dams. *Animal Behaviour*, 54, 509–524.
- Mateo, J., & Holmes, W. (1999). Plasticity of alarm call response development in Belding's ground squirrels (*Spermophilus beldingi*). *Ethology*, 105, 193–206.
- Mitani, J. C., & Brandt, K. L. (1994). Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology*, 96, 233–252.
- Mitani, J., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour*, 135, 1041–1064.
- Molles, L. E., & Vehrencamp, S. (2001). Songbird cheaters pay a retaliation cost: Evidence for auditory convention signals. *Proceedings of the Royal Society, London, B*, 268, 2013–2019.
- Nelson, D., & Marler, P. (1993). Innate recognition of song in white-crowned sparrows: A role in selective vocal learning? *Animal Behaviour*, 46, 806–808.
- Nelson, D., & Marler, P. (1994). Selection-based learning in song development. *Proceedings of the National Academy of Sciences, USA*, 91, 10498–10501.
- Nordby, J. C., Campbell, S. E., & Beecher, M. D. (2007). Selective attrition and individual song development in song sparrows. *Animal Behaviour*, 74, 1413–1418.
- Nordby, J. C., Campbell, S. E., Burt, J. M., & Beecher, M. D. (2000). Social influences during song development in the song sparrow: A laboratory experiment simulating field conditions. *Animal Behaviour*, 59, 1187–1197.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (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., Seyfarth, R. M., & Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cynocephalus ursinus*): Implications for production processes and functions. *Journal of the Acoustical Society of America*, 101, 2951–2963.
- Peake, T. M., Terry, A. M. R., Dabelstein, T., & McGregor, P. K. (2002). Do great tits assess rivals by combining direct experience with information gathered by eavesdropping? *Proceedings of the Royal Society, London, B*, 269, 1925–1929.

- 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.
- Pepperberg, I. M. (2002). *The alex studies: Cognitive and communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.
- Peters, S., Marler, P. R., & Nowicki, S. (1992). Song sparrows learn from limited exposure to song models. *The Condor*, *94*, 1016–1019.
- Pierce, J. (1985). A review of attempts to condition operantly alloprimate vocalizations. *Primates*, *26*, 202–213.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow and Sons.
- Poole, J., Tyack, P., Stoeger-Horwarth, A., & Watwood, S. (2005). Elephants are capable of vocal learning. *Nature*, *434*, 455–456.
- Ramachandran, U., & Coss, R. G. (2000). Age differences in the responses to adult and juvenile alarm calls by bonnet macaques (*Macaca radiata*). *Ethology*, *106*, 131–144.
- Randler, C. (2006). Red squirrels (*Sciurus vulgaris*) respond to alarm calls of Eurasian jays (*Garrulus glandarius*). *Ethology*, *112*, 411–416.
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*, 151–160.
- Riebel, K. (2000). Early exposure leads to repeatable preferences for male song in female zebra finches. *Proceedings of the Royal Society, London, B*, *267*, 2553–2558.
- Russ, J. M., Jones, G., Mackie, I. J., & Racey, P. A. (2004). Interspecific responses to distress calls in bats: A function for convergence in call design? *Animal Behaviour*, *67*, 1005–1014.
- Savage-Rumbaugh, E. S. (1986). *Ape language: From conditioned response to symbol*. New York: Columbia University Press.
- Schmidt, K. A., Lee, Eunice, Ostfield, R. S., & Sieving, K. (2008). Eastern chipmunks increase their perception of predation risk in response to titmouse alarm calls. *Behavioural Ecology*, *119*, 759–763.
- Schusterman, R. J., & Krieger, K. (1986). Artificial language comprehension and size transposition by a California sea lion (*Zalophus californianus*). *Journal of Comparative Psychology*, *100*, 348–355.
- Schusterman, R. J., Reichmuth Kastak, C., & Kastak, D. (2002). The cognitive sea lions: Meaning and memory in the lab and in nature. In M. Bekoff, C. Allen, & G. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 217–228). Cambridge, MA: MIT Press.
- Sewall, K. B., & Hahn, T. P. (2009). Social experience modifies behavioral responsiveness to a preferred vocal signal in red crossbills (*Loxia curvirostra*). *Animal Behaviour*, *77*, 123–128.
- Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, *34*, 1640–1658.
- Seyfarth, R. M., & Cheney, D. L. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour*, *40*, 754–764.
- Seyfarth, R. M., & Cheney, D. L. (1997a). Some general features of vocal development in non-human primates. In M. Husberger & C. T. Snowdon (Eds.), *Social influences on vocal development* (pp. 249–273). Cambridge: Cambridge University Press.
- Seyfarth, R. M., & Cheney, D. L. (1997b). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning and Behavior*, *25*, 249–267.
- Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173.
- Seyfarth, R. M., & Cheney, D. L. (2008). Primate vocal communication. In M. Platt & A. A. Ghazanfar (Eds.), *Primate neuroethology*. Oxford: Oxford University Press.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*, 1070–1094.
- Sherman, P. W. (1977). Nepotism and the evolution of alarm calls. *Science*, *197*, 1246–1253.
- Silk, J. B., Altmann, J., & Alberts, S. C. (2006a). Social relationships among adult female baboons (*Papio cynocephalus*). I: Variation in the strength of social bonds. *Behavioral Ecology & Sociobiology*, *61*, 183–195.
- Silk, J. B., Altmann, J., & Alberts, S. C. (2006b). Social relationships among adult female baboons (*Papio cynocephalus*). II: Variation in the quality and stability of social bonds. *Behavioral Ecology & Sociobiology*, *61*, 197–204.
- Silk, J. B., Seyfarth, R. M., & Cheney, D. L. (1999). The structure of social relationships among female baboons in the Moremi Reserve, Botswana. *Behaviour*, *136*, 679–703.
- Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches: Progress and prospects. *Advances in the Study of Behaviour*, *18*, 1–34.
- Sloan, J. L., & Hare, J. F. (2006). Adult Richardson's ground squirrels (*Spermophilus richardsonii*) ignore rate changes in juvenile alarm calls: Age-differential response urgency perception? *Ethology*, *112*, 896–902.
- Slocumbe, K., Townsend, S. W., & Zuberbuhler, K. (2008). Wild chimpanzees (*Pan troglodytes schweinfurthii*) distinguish between different scream types: Evidence from a playback study. *Animal Cognition*, *12*, 441–449.
- Snowdon, C. T. (1986). Vocal communication. In G. Mitchell & J. Erwin (Eds.), *Comparative primate biology. Behavior, conservation, and ecology* (Vol. 2A, pp. 495–530). New York: Alan R. Liss.
- Spector, D. A., McKim, L. K., & Kroodsma, D. E. (1989). Yellow warblers are able to learn songs and situations in which to use them. *Animal Behaviour*, *38*, 723–725.
- Terrace, H. S. (1979). *Nim*. New York: Knopf.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10*, 121–125.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral & Brain Sciences*, *28*, 675–691.
- Vehrencamp, S. (2001). Is song-type matching a conventional signal of aggressive intentions? *Proceedings of the Royal Society, London, B*, *268*, 1637–1642.
- Watwood, S., Owen, E. C. G., Tyack, P., & Wells, R. (2005). Signature whistle use by temporarily restrained and free-swimming bottlenose dolphins (*Tursiops truncatus*). *Animal Behaviour*, *69*, 1373–1386.
- West, M., King, A. P., White, D. J., Gros-Louis, J., & Freed-Brown, G. (2006). The development of local song preferences in female cowbirds (*Molothrus ater*): Flock living stimulates learning. *Ethology*, *112*, 1095–1107.
- Zuberbuhler, K. (2001). Predator-specific alarm calls in Campbell's guenons. *Behavioral Ecology & Sociobiology*, *50*, 414–422.
- Zuberbuhler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour*, *63*, 293–299.
- Zuberbuhler, K., Cheney, D. L., & Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology*, *113*, 33–42.