Reevaluating Chimpanzee Vocal Signals: Toward a Multimodal Account of the Origins of Human Communication

Adam See

Abstract The vocalizations of chimpanzees have long been thought to be largely genetically predetermined and therefore unlearnable, involuntarily produced, and broadcast indiscriminately. Tomasello (2008) has recently written that, while chimpanzee vocalizations share these constraints and limitations with the vocal displays of all other non-human animals, the attention-getting gestures of chimpanzees are an "evolutionary novelty" because they are, in his estimation, capable of being produced intentionally. As such, chimpanzee gestures are highly significant to discussions of animal cognition and the evolution of human communication. This chapter challenges Tomasello's grounds for restricting this evolutionary novelty to the gestural modality. I argue that, in fact, recent evidence suggests that there is a significant functional difference between certain chimpanzee vocalizations and the vocal displays of other animals and that, based on Tomasello's own criteria for intentionality, gestures do not appear to have a monopoly on intentional communication in chimpanzees. Ultimately, this chapter aims to provide grounds for a multimodal account of the evolution of human communication. I conclude by suggesting that although there is reason to doubt that chimpanzees can communicate intentionally, there is no *more* reason to doubt this ability in the vocal modality than there is in the gestural modality.

Keywords Tomasello • Vocalizations • Intentionality • Chimpanzee • Animal communication • Evolution of language

---

A. See (✉)
Department of Philosophy, Graduate Center of the City University of New York,
365 Fifth Avenue, New York, NY 10016, USA
e-mail: asee@gc.cuny.edu

M. Pina and N. Gontier (eds.), *The Evolution of Social Communication in Primates*,
Interdisciplinary Evolution Research 1, DOI: 10.1007/978-3-319-02669-5_11,

1 Questioning an “Evolutionary Novelty”

It has long been recognized that chimpanzees, along with other great ape species, possess a remarkable capacity for gestural communication. While the vast majority of non-human communicative acts are inextricably bound up with specific emotions, contexts, and environmental cues (Janik and Slater 1997; Cheney and Seyfarth 2010), the past 30 years of work in primatology have provided suggestive evidence that certain gestures of both wild and captive chimpanzees are produced voluntarily and with great circumstantial flexibility [see Pollick and de Waal (2007), Arbib et al. (2008), and Tomasello (2008) for reviews]. Further, much work has been done to substantiate the view that novel gestures are capable of being learned [see Tomasello (1996, 2008) for reviews], socially inherited (Pollick and de Waal 2007), and combined to construct a “simple syntax” (reviewed in Tomasello 2008). Perhaps most impressively, certain chimpanzee gestures appear to be produced “dyadically,” i.e., with sensitivity to the attentional state of the recipient. Liebal et al. (2004) found that, when gesturing to both humans and conspecifics, chimpanzees will reliably exercise the following process: Attempt one gesture, monitor the receiver’s response, and if necessary, walk around the receiver and repeat the gesture or try a different one. As Tomasello (2008: 30) notes, “This shows persistence to a goal with adjusted means as necessary—the prototype of intentional action.” The fact that chimpanzees appear to employ “practical reasoning” in gestural communication strongly suggests that they possess a theory of mind, i.e., that they attribute mental states such as attention to others (Tomasello 2008; Premack and Woodruff 1978).[1] Chimpanzee gestures are, therefore, highly significant to discussions about animal minds and the evolution of human communication. According to Tomasello, “attention to the attention of the other during communication is unprecedented in non-primate, and maybe even non-ape, communication” (33).

It is because of this fact that Tomasello (2008) draws a “sharp contrast” between the attention-getting gestures of chimpanzees (which he calls “intentional signals”), and the mere “communicative displays” that encompass all other acts of animal communication. Setting aside any contentious terminological issues for the moment, Tomasello’s view is that there is no *functional* difference between deer horns, peacock tails, bee dances, teeth-bearing, and, salient to this chapter, all animal vocalizations including those of non-human primates, song birds, and cetacea. Tomasello’s claim that the gestures of chimpanzees are an “evolutionary novelty” is therefore a very strong one. It is also, perhaps surprisingly, not very

---

[1] Since Premack and Woodruff’s original discussion, “theory of mind” has been used to describe and/or explain an extensive range of phenomena. In the interest of clarity, I purport here to use the term exactly as Tomasello (2008) does in its relationship to what he calls “intentional signals.” For a discussion of Tomasello’s understanding of theory of mind and its significance to intentional signals see Sect. 2 below.

controversial in the modern literature. The view that chimpanzee vocalizations are, for the most part, unlearnable, inflexibly tied to emotions, involuntarily produced, genetically predetermined, and broadcast indiscriminately, is in fact widely accepted [see Arbib et al. (2008), Hammerschmidt and Fischer (2008), and Cheney and Seyfarth (2010) for reviews].

There are, however, studies as recent as this year that support a contrary position on the communicative potential of chimpanzee vocalizations. In response to Tomasello's view, I argue here that there *is* a significant functional difference between certain vocalizations of chimpanzees and the vocal displays of other animals. Gestures do not have a monopoly on intentional communication in non-human primates (henceforth primates). This chapter suggests both that certain vocalizations of chimpanzees satisfy all of Tomasello's criteria for intentional signals and that any skepticism pertaining to the intentionality of these vocalizations applies equally to their gestural counterparts. If this claim can be substantiated, Tomasello's (2008) argument that the origins of human communication emerged primarily from primate gestures may demand revision.

The most common argument against the evolutionary relevance of the great ape vocal modality is a physiological one [see Hammerschmidt and Fischer (2008) for a review]. Like all non-human animals, great apes have very limited flexibility in vocal production (Cheney and Seyfarth 2010). So although chimpanzees have, for instance, been shown to socially inherit novel whistling behavior (Crockford et al. 2004) and flexibly employ goal-oriented "raspberry" lip-purses to achieve communicative ends (Leavens et al. 2004; Russell et al. 2005; Hostetter et al. 2007), in this chapter I follow Hopkins et al. (2007) in distinguishing "vocalizations" from mere "sounds" such as lip-smacks and whistles and restrict the evidence considered to vocal signals generated with use of the vocal cords.

The general trajectory of this chapter proceeds as follows. Since Tomasello (2008) is not explicit about the criteria he uses to distinguish "communicative displays" from "intentional signals," I begin by deriving three general criteria from his argument. These are the presence of (1) social and asocial learning in the ontogenetic development of signaling behavior, and (2) contextual flexibility and (3) attention to the attention of the receiver during signal production. I then provide what I hope to be compelling evidence that great apes may be capable of vocalizations that meet all of these criteria and thus should, by Tomasello's own account, be classified as intentional signals that differ significantly from common vocal displays.

It is imperative to note that my argument here is *not* that the vocal signals of chimpanzees are as flexible or even as evolutionarily significant as chimpanzee gestures, but rather that Tomasello's arguments against their communicative potential are unjustified. To this end, I conclude by suggesting that although there is reason to doubt that chimpanzees can communicate intentionally, the fact that they utilize attention-getting (henceforth AG-) vocalizations in ways that are functionally equivalent to AG-gestures demonstrates that there is *no more* reason to doubt intentionally in the vocal modality than in the gestural modality.

2 "Communicative Displays" and "Intentional Signals"

Tomasello (2008) unfortunately does not list any specific criteria that distinguish "communicative displays" from "intentional signals." Rather, he tends toward defining gestural signals negatively in relation to what he calls "vocal displays." This section aims to identify and make explicit the criteria that appear to be tacitly operating in his argument in order to apply them in subsequent sections.

Defined as broadly as possible, a signal is any sign or indication of a given state of affairs. In the animal kingdom, pelt coloration, nest construction style, beak size, vocalizations, and gestures are all signals because they all convey information. There is therefore a very important sense in which all animal displays simply *are* signals. That said, discussions of the evolution of human communication will inevitably demand more precise terminology. This is because intentionality, voluntarism, and developmental and contextual flexibility now become highly salient traits in the psychology of both senders and receivers. Tomasello largely avoids the task of psychologically carving up the world of animal communication by drawing a very thick line in the sand. Any signal that is not produced *intentionally* is a display:

> Communicative displays are prototypically physical characteristics that in some way affect the behavior of others, such as large horns which deter competitors or bright colors which attract mates. Functionally, we may group with displays reflexive behaviors that are invariably evoked by particular stimuli or emotional states and over which the individual has no voluntary control. Such inflexible physical and behavioral displays, created and controlled by evolutionary processes, characterize the vast majority of communication in the biological world (14).

In "sharp contrast," Tomasello writes that "intentional signals"…

> …are chosen and produced by individual organisms flexibly and strategically for particular social goals, adjusted in various ways for particular circumstances. These signals are *intentional* in the sense that the individual controls their use flexibly toward the goal of influencing others (14).

It is worth highlighting here that Tomasello—presumably for clarity—is using the word "signal" to demarcate volitional or intentional modes of communication, while "display" is used exclusively in reference to communicative signs that the signaler has no control over. I will stick to Tomasello's terminology in what follows. Though Tomasello does not explicitly classify primate vocalizations as displays, textual evidence that he classifies them as such is evident in the following "(many gestures) are as genetically fixed and inflexible as primate vocalizations—and thus should be called displays—an important subset are individually learned and flexibly used, especially in the great apes, and so may be properly called intentional signals" (20). It is clear then that while displays and intentional signals both share the function of influencing the behavior of others, the latter possess two inextricable qualities: *volitional flexibility,* i.e., "the individual controls their use," and *agent-directed behavior* with "the goal of influencing others." Note that for Tomasello a theory of mind is implicit in the latter. In order to intentionally alter "the attentional state of the recipient, (…) the communicator needs some kind of cognitive model of how the recipient perceives the signal and acts as a result" (45). It is therefore evident that for Tomasello, the production of intentional signals involves second-order intentionality (Dennett 1987). Namely, in order for their signals to modify not just the behavior, but also the mental states of others, the signaler must possess an understanding of both their own mental states and those of others, e.g., alarm calls are intended to alter an ignorant receiver's *knowledge* of the situation, and cause them to flee because of this new information. Second-order intentionality is contrasted with first-order intentionality, where signals are produced with a desire to influence the behavior of others but, crucially, the signaler need not understand the mind of the receiver in order to predict and recognize the behavioral effects of their signal, e.g., the receiver flees. Tomasello (2008) appears to suggest that chimpanzee vocalizations, like all "communicative displays," are produced with *zero-order* intentionality; namely, the signaling behavior is purely reflexive and is involuntarily produced with zero mentality attributed to the receiver(s). Since chimpanzee gestures are often produced with persistence and attention to the attention of the receiver—behavior that best exemplifies second-order intentionality—Tomasello uses behavior of this nature as his primary criterion for intentional signals.

On a more foundational level, Tomasello (2008: 21) identifies a strong connection between volitional flexibility and advanced capacities for signal learning. Whether or not a given behavioral trait requires learning can be a strong indication as to whether that trait is genetically determined, and as a result, the extent to which its production may be voluntary or intentional. Tomasello thus uses the presence of *social and/or asocial learning* in gestural communication, and the apparent lack of this characteristic in the chimpanzee vocal modality, as further evidence for the lack of intentionality in the latter.

It is clear then that if a signal is capable of being learned and utilized flexibly with persistence and attention to the attention of the other, that signal satisfies Tomasello's criteria for intentional production. As mentioned above, this chapter confronts these criteria beginning with social learning, continuing on to communicative flexibility, and finally engaging intentional production. It is worth stressing that Tomasello denies that the vocalizations of great apes satisfy any of these criteria. According to Tomasello, the modality of ape gestures "contrasts totally with their unlearned, inflexible, and emotional vocalizations indiscriminately broadcast to the world" (320).

3 Social and Asocial Learning

In *Origins of Human Communication*, Tomasello spends a considerable amount of time providing evidence that chimpanzees are capable of learning novel gestures—a feat unavailable to the vocal modality. This is presumably because, according to Hammerschmidt and Fischer (2008: 93), a "prerequisite for a high degree of (communicative) flexibility is learning, in terms of both production and comprehension." The capacity for social and/or asocial learning is therefore a logical precondition for both flexibility and intentionality in primate vocalizations. The overarching goals of this section are to demonstrate that the existing evidence in support of vocal learning in chimpanzees is (1) comparable to the evidence in support of learning in chimpanzee gestures, and (2) currently insufficient to justify Tomasello's conclusions on the matter.

The parameters of which acts of behavioral transmission should be classified as instances of "social learning" have been the subject of considerable debate. As Galef (1976) observed, interactions among conspecifics in several species are known to affect the acquisition and expression of complex behaviors. There is clearly a difference, however, between a social influence on the *use* or *application* of "innate" behaviors and behavior that either is *itself* modified or that would not exist if the subject were not socialized to exhibit it. In their oft-cited paper on social learning in animals, Janik and Slater (1997) usefully distinguish between "contextual learning" and "production learning." The former refers to learned modifications in the contextual usage of the signal, and the latter "refers to instances where the signals themselves are modified in form as a result of experience with those of other individuals." *Vocal learning*, they argue, "is defined by production learning in the vocal domain." In order for a vocalization to be classified as "socially learned," then, I submit that it must either be (1) a novel vocalization that is only used in particular groups of conspecifics, (2) a vocalization that individuals do not develop when raised in isolation, or (3) an idiosyncratic vocalization shared primarily by the mother and her own offspring. While species that meet any or all of these criteria are extremely rare [see Cheney and Seyfarth (2010) for comments], the well-documented case of the zebra finch has, for instance, demonstrated that production learning of vocalizations does occur outside of the human domain. The juvenile zebra finch will learn the complex, idiosyncratic calls of its mother during its first few weeks and then will begin to lose them with age as its hearing deteriorates (Fehér et al. 2009). The zebra finch is thus a paradigm case of vocal production learning in the animal kingdom.

Though he is not explicit about this, Tomasello appears to use "production learning" as his criterion for social learning. Support for this claim is may be identified in Tomasello's two major arguments about the relative uniqueness of chimpanzee gestures. First, "individuals with significant human contact invent or learn different kinds of novel gestures quite easily," and second, "there are many and very large individual differences in the gestural repertoires of different individuals of the same species, even within the same group, including some idiosyncratic gestures produced by single individuals" (21). Tomasello then contrasts these points with claims that "within any monkey and ape species all individuals have the same basic vocal repertoire, with essentially no individual differences in repertoire," and that "monkeys raised in social isolation and monkeys cross-fostered by another species (…) still produce their same basic species-typical vocalizations" (16).

I submit that Tomasello's first claim is false and that his second claim is deceptively used and largely irrelevant to his argument. I will confront these two claims in reverse-order beginning with the latter. In their comprehensive overview of research in primate vocalizations, Hammerschmidt and Fischer (2008: 94) observe that "Most of the evidence accumulated about vocal development comes from studies on monkeys, while little is known about the vocal development of apes." Tomasello himself actually recognizes this strong asymmetry in the available data (Tomasello and Zuberbuhler 2002). While it is true that monkey vocalizations almost certainly do not demonstrate production learning—e.g., squirrel monkeys deafened at birth acquire structurally equivalent calls to normal monkeys (Winter et al. 1973)—notably, the *same is true of monkey gestures,* which are largely identical across species (Arbib et al. 2008). Indeed, the general consensus among primatologists is that monkey and ape species differ significantly in cognitive capacity and communicative modality [see Arbib et al. (2008) for a review]. Tomasello's use of the limitations of monkey vocal development is therefore not sufficient to support his claim that *ape* vocalizations are "unlearned" or "not individually learned" (Tomasello 2008: 320, 33). This claim brings me to my second point, which is that though they are few in number, there do exist studies that suggest a marked capacity for social and asocial vocal production learning in chimpanzees.

Tomasello's claim that within ape species "all individuals have the same basic vocal repertoire" is challenged on a number of levels. First, Leavens et al. (2004) demonstrated that captive chimpanzees produce "novel" grunts that vary in tone, timbre, and length previously unheard in the wild.[2] A number of studies since have corroborated Leavens et al.'s interpretation that these vocalizations appear to be used exclusively among captive chimpanzees toward humans for "attention-getting" purposes (Russell et al. 2005; Hostetter et al. 2007; Hopkins et al. 2007; Taglialatela et al. 2012). Until recently, primatologists have been in the dark as to potential causal explanations for the emergence of these signals unique to apes raised in captivity. One plausible hypothesis is offered by Leavens et al. (2010), who argue that some captive apes learn, asocially, how to solve a problem unique to their environment. In the wild, chimpanzees can attain food whenever they want; only in captivity do they often need to *ask* for it. It is therefore possible that captive chimps are spontaneously learning to use AG-vocalizations (and humans) as tools to achieve desired ends outside their natural environment.

In a recent paper entitled "Social learning of a communicative signal in captive chimpanzees," Taglialatela et al. (2012) offer data in support of social transmission of AG-vocalizations. This study found that juveniles raised by their biological mothers in captivity are far more likely to exhibit these novel vocalizations than those juveniles raised by humans in a nursery environment. Further, a strong correlation exists

---

[2] Note that I am referring here to what are commonly referred to as "extended food grunts" and not "raspberry" sounds since only the former employ the vocal chords. Though raspberry sounds have never been observed in the wild and serve the same AG-function as novel grunts, they involve only a pursing of the lips. The appropriate place of whistles and lip-smacks in discussions of the evolution of human communication is certainly interesting yet remains beyond the scope of this chapter.

between *which* of the two basic AG-vocalizations the mother typically uses and the one picked up and utilized by the infant. From these discoveries the authors conclude that, "These data support the hypothesis that social learning plays a role in the acquisition and use of communicative vocal signals in chimpanzees" (3). Captive chimpanzees therefore exhibit novel vocalizations previously unheard among conspecifics in the wild and there is a strong correlation between idiosyncratic vocal production in mothers and their own offspring. As evidence for vocal production learning in captive chimpanzees, these studies challenge Tomasello's claim that while ape gestures are "individually learned," "this is not true of ape vocalizations" (33).

There have also been studies that suggest production learning in the vocalizations of wild chimpanzees. Using spectrographic analyses in two separate studies, Mitani et al. (1999) and Mitani and Nishida (1993) found significant differences in the vocal calls of adjacent and distal chimpanzee groups in East Africa. Crockford et al. (2004), moreover, found that the pant hoots of male chimpanzees living in three adjacent communities along the Ivory Coast differ more strongly from each other than either of them do from those of a community over 70 km away. They conclude that since neither habitat nor genetics appear to account for this difference, these chimpanzees seem to have "actively modified" their pant hoot structure to better facilitate in-group identification. While "actively modified" is surely a contentious interpretation of the data, note that even the more modest interpretation, i.e., unconscious "call convergence," still demonstrates the learning and dissemination of idiosyncratic call *structures* (Marshall et al. 1999). Marler (1991), for instance, has suggested that "action-based learning" can selectively reinforce structural call variants in chimpanzees. Regardless of how one interprets this data, these findings are clearly more suggestive of Janik and Slater's notion of *production learning* than they are of *contextual learning*.

Since the literature on this subject is rare and still in its nascent stages, Tomasello is unjustified in his conclusion that while ape gestures are "individually learned," "this is not true of ape vocalizations" (33). Contrary to Tomasello's view, chimpanzees appear to be capable of socially acquiring vocalizations as per the stringent criteria of Janik and Slater (1997) *and* in comparison with Tomasello's data on ape gestures. For instance, Tomasello's claim that apes with significant human contact learn "different kinds of novel gestures" applies equally to vocalizations. Also, his claim that various chimpanzee groups exhibit different gestural repertoires is also evident in chimpanzee vocalizations. As I mentioned above, this is not to suggest that novel vocalizations are as easily and flexibly acquired as gestures, but rather that the production of gestures and vocalizations do not differ in their general capacity to be learned by chimpanzees.

4 Contextual Flexibility and Smart Receivers

As mentioned above, Tomasello (2008) never offers a positive definition of volitional flexibility in animal minds. Instead, he tends to define volition negatively in relation to inflexible "vocal displays" (16). This tactic is slightly misleading since a negative definition is hardly a definition at all. It is, however, understandable. As Povinelli and collaborators (Povinelli and Eddy (1996), Povinelli and Vonk (2006)) have repeatedly pointed out, studies of the workings of animal minds are almost exclusively limited to interpretations of animal behavior. While the literature is slowly incorporating more work in neuroscience, Tomasello does not present any evidence suggesting that chimpanzees lack neurological capacities for "volitional" communication. Tomasello's means of defining volitional flexibility is therefore to contrast various communicative *behaviors*: A "signaler has intentional control over the signal" (13) rather than being "controlled by evolutionary processes" (14) if the signal is not "ritualized" or accompanied by an external stimulus or an emotion (16–17). As such, he claims, AG-gestures are ideal candidates for volitional flexibility because they occur in a wide variety of contexts, most of which are non-urgent and divorced from impulsive emotions. Tomasello thus argues that while vocalizations have an evolutionary history of association with urgent functions such as alarm calls and mediating conflict, chimpanzees occasionally use AG-gestures to call attention to displays that initiate common, everyday activities such as sex, play, nursing, begging, and grooming (20). Tomasello's principle criterion for volitional flexibility is therefore the capacity to use a signal in contexts that are not ritualized or emotionally urgent.

Tomasello understands the apparent contextual limitations of chimpanzee vocalizations and the non-emotional expression of these vocalizations to be intimately related: "the connection between a vocal call and its eliciting emotion or situation is mostly very tightly fixed; non-human primates do not vocalize flexibly by adjusting to the communicative situation" (16–17). For concrete evidence of this claim, Tomasello again relies almost exclusively on monkey data (18–20). For reasons discussed in the previous section, this move is ultimately inadequate to make his point. The only additional evidence that Tomasello provides with respect to chimpanzees is a field observation made by Jane Goodall in the 1980s that "The production of a sound in the *absence* of the appropriate emotional state seems to be an almost impossible task for a chimpanzee" (17) and the fact that chimpanzees make pant hoot calls in the presence of food even when everyone else is already present. However, as the authors of this latter study (Clark and Wrangham 1994) suggest these pant hoots are more likely signals of *status* than alarm calls. Tomasello is therefore unjustified in using this example alongside those of alarm calls to (surreptitiously) suggest the emotional inflexibility of chimpanzee vocalizations.

These two sources are the extent of Tomasello's evidence with respect to chimpanzees. The rest of his evidence is taken from studies on monkey alarm calls wherein, according to Seyfarth and Cheney (2003: 168), "Listeners acquire information from signalers who do not, in the human sense, intend to provide it." In these cases, "alarms" appear to be involuntary, emotional vocalizations produced reflexively in response to stimuli that nearby animals merely "eavesdrop" on (Cheney and Seyfarth 2005, 2007). Monkey alarm calls are therefore unlikely to be flexibly or intentionally produced. Macaque mothers, for instance, do not vocalize when predators approach their young so long as they are themselves a safe distance away (Cheney and Seyfarth 2005).

Both Cheney and Seyfarth (2005) and Tomasello (2008) explain this surprising behavior by suggesting the absence of a theory of mind in macaques. Presumably, if the mother understood her juvenile's *perspective,* i.e., as being *unaware* of the predator, she would have informed her of its presence (Cheney and Seyfarth 2005). Regardless of the strength of this reasoning—or, for that matter, whether or not certain monkey species possess a theory of mind[3]—Tomasello's argument here draws attention to another serious flaw in his reliance on monkey data to make claims about apes: If one does not believe that monkeys have a theory of mind, then it is unfair to use them as evidence for the lack of cognitive capacities in a species that one believes *does* possess a theory of mind. This criticism is particularly damning for Tomasello's argument that chimpanzee vocalizations are incapable of intentional production—a capacity, recall, that Tomasello claims requires a theory of mind. Notably, there do not appear to be any available studies that demonstrate analogous behavior with respect to chimpanzee alarm calls. If there is, such studies do not appear in any of the major review papers on the subject, i.e., Cheney and Seyfarth (2010), Hammerschimdt and Fischer (2008), or Arbib et al. (2008).

However, even if such evidence did exist its utility would ultimately be circular. This is because alarm calls are *exclusively* sounded in urgent, emotionally charged contexts. Another significant flaw in Tomasello's argument that chimpanzee vocalizations are inextricable from emotions is therefore that the only examples he provides involve contexts where this must be the case. For evidence of vocalizations in non-urgent contexts, I refer back to the attention-getting vocalizations described in the previous section. Like ape gestures, the AG-vocalizations used exclusively with humans have been utilized in non-predatory and non-competitive contexts without direct emotional stimulation. In fact, Russell et al. (2005) and Hopkins et al. (2007) conducted experiments to control for precisely this factor. Whereas Leavens et al.'s original (2004) study used visible food to elicit AG-vocalizations, Russell's team found that chimpanzees will utilize AG-vocalizations to solicit a necessary tool from a human in order to attain food that is out of sight. This finding suggests not only that these vocalizations are not automatic, emotional reactions to stimuli (food), but also that they are used selectively to communicate with others in instances of practical reasoning. According to Tomasello, "practical reasoning" about others' perspectives, e.g., a theory of mind, "underlies flexible communication" and is fundamental to intentional communication (48-9).

Further support for the claim that chimpanzees are capable of selectively using AG-vocalizations may be found in the results of Hopkins' group. Hopkins et al. (2007) recorded the frequencies of both traditional "food" vocalizations and AG-vocalizations when the chimpanzees were presented with the food alone, the human alone, or the food visible with a human. They found that chimpanzees produce significantly more "food" vocalizations with food alone than in the other

---

[3] See Flombaum and Santos (2005) for evidence suggesting that rhesus macaques can be sensitive to the attention of others and engage in practical reasoning about others' perspectives. Further, Wich and de Vries (2006) offer evidence that Thomas langur monkeys possess the capacity for goal-directed vocal production.

two conditions and significantly more AG-vocalizations when food was presented along with a human. Hopkins et al. conclude that "chimpanzees recognize functional differences" between various calls and can "selectively produce" AG-vocalizations in particular contexts.

Recent studies of chimpanzees in their natural habitat corroborate these findings, further challenging Tomasello's claim that chimpanzee vocalizations are "broadcast indiscriminately." In line with the studies on captive chimpanzees, wild chimpanzees have been documented producing agent-directed vocalizations to conspecifics [Goodall (1986), Townsend et al. (2008), Slocombe et al. (2010), Laporte and Zuberbühler (2010), Schel et al. (2013a, 2013b); also see Halloran (2012) for spectrographic analyses of context-specific, agent-directed AG-vocalizations among conspecifics in captivity]. Contrary to Tomasello's questionable interpretation of Clark and Wrangham's (1994) findings on chimpanzee food calls, a recent study by Slocombe et al. (2010) suggests that wild male chimpanzees "produce food-associated calls selectively in the presence of important social partners." The authors claim that the stable, long-term relationships formed by male chimpanzees play an important role in whether or not food vocalizations are produced. Further, Schel et al. (2013b) have more recently provided evidence of food calls by wild male chimpanzees being directed at specific individuals based on rank and friendship. They conclude by suggesting that "chimpanzee food calls are not simply reflexive responses to food, but can be selectively directed at socially important individuals." They continue: "Our findings are thus inconsistent with traditional views of primate vocalizations as inflexibly and indiscriminately produced. Instead, our results indicate that great apes can produce semantically meaningful calls in a highly selective, recipient-directed manner." Audience composition therefore appears to have a marked influence on not only *when* male chimpanzees produce food vocalizations, but also to *whom* these calls are directed.

Laporte and Zuberbühler (2010) likewise provide field data suggesting that female chimpanzees "selectively deploy vocal signals depending upon the social context." The authors reported that female chimpanzees are more likely to produce agent-directed vocalizations toward lower-ranked males in the absence of the group's alpha male. Complementary findings can be found in an earlier paper entitled "Female Chimpanzees Use Copulation Calls Flexibly to Prevent Social Competition," wherein Townsend et al. (2008) report that female chimpanzees called significantly more around high-ranking males and suppressed vocalizations if high-ranking females were within earshot. According to the authors, since competition among females is dangerously high in the wild, "Copulation calling may therefore be one potential strategy employed by female chimpanzees to advertise receptivity to high-ranked males, confuse paternity and secure future support from these socially important individuals." Townsend et al. claim that their findings contradict an established hypothesis that copulation calls are merely a sexually selected trait to instigate male–male competition. Females in fact appear to have considerable selective control over when they vocalize and whom they vocalize to—an hypothesis reinforced by a hormone analysis that demonstrated that their subjects vocalized at times unrelated to their fertile period.

These results are further corroborated by studies measuring the effects of audience composition on both the selection and structure of chimpanzee vocalizations during inter-group conflict (Wilson et al. 2001; Slocombe and Zuberbühler 2007). Slocombe and Zuberbühler (2007), for instance, found that victims of agonistic interactions "alter the acoustic structure of their screams depending on the severity of aggression experienced, providing nearby listeners with important cues about the nature of the attack," and that if one audience member matched or surpassed the aggressor in rank, these victims reliably "exaggerated the true level of aggression experienced." The authors therefore conclude that "chimpanzees possess sophisticated understanding of third-party relationships, so-called triadic awareness, and that this knowledge influences their vocal production." Significantly, this data suggests that Tomasello is incorrect that the fact that chimpanzees vocalize most often in emotional contexts necessarily reflects negatively on the agent-directedness and contextual flexibility of their calls.

Notably, all of the studies discussed in this section suggest that vocalizations are used to facilitate intentional, goal-directed behavior. As will become relevant in the next and final section, this behavior can also be explained by employing learned behavioral rules rather than, as it might initially appear, a theory of mind. As I aim to show, however, the same criticism applies to analogous studies of chimpanzee gestures.

Before turning to this issue, it is important to note that the studies discussed thus far also suggest that a limited call repertoire can nonetheless demonstrate flexibly in function when semantic context is taken into account. In his analysis of primate vocalizations, Tomasello restricts his discussion to the vocal capacities of the sender, but once context is taken into consideration, it becomes clear that one does a great disservice to the communicative potential of not only primates, but all vocalizing species, to discount the role of the receiver in information transmission. The context in which vocalizations are produced can provide "smart receivers" with a wealth of salient information. In their survey of animal vocal communication, Cheney and Seyfarth (2010: 93) claim that even though vocal production is largely restricted across the animal kingdom, one should not assume that these physiological limitations correspond to limitations in signal *content:*

> 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.

Similarly, Hammerschmidt and Fischer (2008: 113) conclude their (primarily negative) summary of vocal limitations in primates with the following: "It might in fact be the case that smart listeners lift some of the pressure off senders because they not only are able to perceive fine-grained differences among calls but also are able to make rich interpretations of calls—in other words, to attribute meanings to them." Because primate species typically live in large communities where the social hierarchy is strictly observed and—as in the case of chimpanzees—often shifting, the *context* of a given vocalization might be thought to carry just as much or even more information than the *type* of vocalization produced. Playback experiments by Cheney and Seyfarth (1990, 2005, 2007) show that baboons recognize specific breaks in social hierarchy and therefore deduce complex "social narratives" from calls. This is a significant feat when one considers that baboon troops live in communities as large as 150 members. The fact that monkeys have the cognitive capacity to identify the specific vocalizations of individual group members has also been demonstrated in ape species [see Berlin (2006) for a review]. Kojima et al. (2003), for instance, found that captive chimpanzees can not only successfully match the pant hoots, grunts, and screams of conspecifics to their photographs, but can also identify both individuals during "duets." Further, Levréro and Mathevon (2013) identified distinctive vocal "signatures" in adult and infant chimpanzees that were consistently present in a variety of different types of call from the same individual. The authors suggest that chimpanzee vocal signatures have tremendous evolutionary value as they, among other functions, facilitate in-group identification and aid mothers in keeping track of their offspring.

In response to their findings on baboons, Cheney and Seyfarth (2005: 149–150) express their surprise that "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." While this may be true with monkey species, there do exist studies demonstrating the use of "creative call combinations" by chimpanzees (Crockford and Boesch 2003, 2005). For instance, the same year that Cheney and Seyfarth made this claim, Crockford and Boesch (2005: 397) conducted a spectrographic analysis of calls made by chimpanzees in the Tai Forest finding that over half of their vocalizations "occurred in combination with other vocalizations or with drumming," and that overall, these chimpanzees utilize a total of "88 different types of combinations" each produced in "specific contexts." They therefore conclude that vocal call combinations are an "important" part of chimpanzee communication that "increase message complexity" by increasing "the range of information that can be decoded by listeners."

Both the controlled and field studies discussed in this section strongly suggest that chimpanzees not only have voluntary control over their ability to vocalize, but also *which* vocalizations they use and *who* they vocalize to. We may conclude, then, that Goodall's field observation that vocalizing in the absence of emotions appears to be an "impossible" task for a chimpanzee is false. Further, when combined with the notion of smart receivers and the extensive use of call combinations, chimpanzee vocalizations have impressive potential for contextual flexibility *even when* produced in emotional contexts. It is therefore reasonable to conclude that these vocalizations meet Tomasello's second criterion for contextual flexibility in both signal production and utility. While this is certainly significant, the criterion that truly separates "intentional signals" from common displays is the third and final criterion: that apes vocalize referentially and discriminately, i.e., with attention to the attentional state of the receiver.

5 Intentionality, Theory of Mind, and the Logical Problem

Though certainly significant in its own right, the evidence and argumentation produced thus far have been intended as foundational to a considerably larger project toward which much future work needs to be done. By attempting to substantiate the hypothesis that chimpanzee vocalizations have the potential to be produced intentionally with volitional flexibility, I have taken a bottom-up approach to suggesting that certain types of their vocalizations are socially learned (and thus are not genetically predetermined) and can be flexibly and selectively produced in non-emotional and non-ritualized contexts. Although Tomasello denies chimpanzee vocalizations even these more basic qualities, they are not terribly rare among vertebrates in the animal kingdom (Cheney and Seyfarth 2010). What remains to be demonstrated is that apes possess the extraordinary ability of using their vocalizations *intentionally*—a capacity that, Tomasello claims, outside human communication only ape gestures appear to facilitate.

In the previous section, I cited studies suggesting that chimpanzees are capable of "selectively choosing" their vocalizations in interactions with humans and conspecifics and that they use novel "attention-getting" vocalizations specifically in these contexts. What is unclear is whether they have learned new behavioral rules, i.e., use vocalization *A* in situation *B* to receive a given reward, or whether they are choosing particular vocal signals intentionally with the goal of altering not just the *behavior* (first-order intentionality) but the mental state, i.e., the "attention" (second-order intentionality), of the human and therefore possess a theory of mind.

It is important to stress here that these two explanations, i.e., theory of mind and behavioral rules, are "functionally equivalent" (Cheney and Seyfarth 2005: 138). In other words, they both achieve the same result, which, in the controlled studies above, is receiving food. This fact has led to perhaps the most highly debated issue in the animal minds literature, which is often referred to as the logical problem. The logical problem states that since all we can observe is an animal's behavior, there is no conceivable way to distinguish whether that animal is employing a theory of mind or a behavioral rule when interacting with others [see Povinelli and Vonk (2006) and Lurz (2011) for comments]. Because there is no functional difference between theory of mind and behavioral rules, and because the latter are far less cognitively taxing, it is scientifically irresponsible to assume that non-human animals possess a theory of mind (Povinelli and Vonk 2006). In this section, I remain neutral on the debate itself and argue that the force of the logical problem applies equally to claims of second-order intentionality in both vocal and gestural signals.

I will begin with a brief overview of Tomasello's argument which again relies almost exclusively on data taken from monkey species in the context of alarm calls. According to Tomasello (2008), since primate vocalizations are "broadcast indiscriminately to everyone nearby," this means that "psychologically" the signaler "need not pay any attention to the recipients, and indeed cannot easily direct vocal calls to selected individuals to the exclusion of others" (18). Elsewhere he claims that, besides vocalizing in slightly modified forms to distinguish different amounts of food, “great apes do not produce any referentially specific calls” (16). Tomasello’s argument for intentional communication in ape gestures ultimately boils down to the fact that, when gesturing, apes occasionally pay attention to the attention of the receiver. Chimpanzees are known to alternate their gaze between food and human subjects while gesturing (Leavens and Hopkins 1998) and, monitoring the receiver’s response, “repair” communication when it has failed by moving closer and/or trying different gestures (Liebal et al. 2004; Leavens et al. 2005). These studies, Tomasello notes, “show persistence to a goal with adjusted means as necessary—the prototype of intentional action.”

What Tomasello (2008) conspicuously does not mention is that the cited Leavens and Hopkins study notes that both gestures *and* vocalizations were utilized during gaze alteration. Indeed, they conclude their paper as follows:

> This gesturing or vocalizing while alternating the gaze between the communicative interactant and the object of interest cardinally represents what has been called in both the human development literature and the literature pertaining to communication in apes, intentional communication (819).

Tomasello’s selective emphasis is also demonstrated in his reference to the Leavens et al. (2004) study, which found that chimpanzees frequently use their AG-vocalizations as *alternatives* to physical readjustment in situations when an inattentive human is facing the wrong direction. In fact, several studies found that chimpanzees modulate their vocal production depending on the orientation (Hostetter et al. 2001) and visual attention of the human receiver (Krause and Fouts 1997; Theall and Povinelli 1999; Bodamer and Gardner 2002; Hostetter et al. 2007; Hopkins et al. 2007). Two of these studies (Bodamer and Gardner 2002; Leavens et al. 2004) found that when one vocalization failed to catch the attention of a human, chimpanzees were *more likely* to try a different vocalization than when the original AG-vocalization was successful. In their analysis of data from nine studies on the use of AG-vocalizations, Hopkins et al. (2007) conclude that the general findings suggest that chimpanzees differentially produce vocalizations according to different attentional cues. The fact that all of these studies were released prior to *Origins of Human Communication* makes it surprising that Tomasello does not at least acknowledge them.

In the human development literature, tactics of persistence and elaboration have frequently been noted as indicators of intentional communication (Bates et al. 1975; Golinkoff 1993). We have seen that Tomasello also clearly classifies signals that demonstrate these communicative tactics as being indicative of intentional production. In line with this literature, Leavens et al. (2005, 2010) found that chimpanzees used tactics of persistence and elaboration to vary “their signals within a modality that was appropriate to the attentional status of a human” to “rapidly accommodate changes” in the human’s attention. In a study of 110 chimpanzees, Leavens et al. (2010) found that both their AG-gestures and AG-vocalizations “followed a logical and efficient pattern of modality-specific permutations” as the chimpanzees continued “to elaborate in a tactically efficient manner throughout a minute-long episode in which their communication was having no apparent effect." The results of these studies certainly suggest, in Tomasello's words, "persistence to a goal with adjusted means as necessary."

Not only do all of these studies suggest that chimpanzee vocalizations can be produced, suppressed, and elaborated upon depending on the attentiveness of a human receiver but, more significantly, that ape vocalizations can, like gestures, be produced intentionally. This is, at least, the conclusion that one is committed to if, like Tomasello, one is convinced that when gestures are used in the same context they exhibit "the prototype of intentional action."

This is precisely the conclusion that Schel et al. (2013a) arrive at in their recent study on wild chimpanzees entitled "Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality" which assessed whether chimpanzees recognize mental states of knowledge and ignorance in conspecifics and then utilize that information selectively in their alarm calls. Schell et al. claim that "In contrast to gestural research, intentionality has rarely been the focus of primate vocal research, thus it is vital that directly comparable evidence is gained to empirically test whether great ape vocal production engages first-order intentionality." Their experiment elaborated upon findings published the previous year by Crockford et al. (2012). Crockford's group ran a field study that demonstrated that chimpanzees were more likely to produce alarm calls in response to a predator when in the presence of unaware or ignorant group members then they were when in the presence of group members with clear knowledge of the threat. In this experiment, a model viper was placed on the projected travel path of a group of 33 chimpanzees, whose vocalizations were then recorded. The authors found that "alarm calls were significantly more common if the caller was with group members who had either not seen the snake or had not been present when alarm calls were emitted," thus concluding that "chimpanzees monitor the information available to other chimpanzees and control vocal production to selectively inform them." Though the results of this study are strongly suggestive of intentional communication, the experiment suffers from one potential oversight: in each iteration, the signaler was also ignorant of the predator, so although *more* calls were given in the presence of unaware group members, the signaler's initial calls may have been produced automatically upon discovery of the snake. This complementary explanation would suggest that the signalers original intention was not to alert group members after all but was rather a reflexive expression of their own fear, i.e., zero-order intentionality.

In their own version of this experiment, Schel et al. (2013a) corrected for this oversight and in doing so "tested the production of chimpanzee vocal signals across multiple markers of intentionality, in a comparable manner to chimpanzee gestures," a task never attempted in previous studies. They begin by distinguishing three distinct types of alarm calls: "soft huus" (SH), "alarm huus" (AH), and "waa barks" (WB). Upon encountering the snake alone in the absence of other group members, subjects only elicited the comparably low, short-ranged SH vocalizations. The other two, much louder and abrasive alarm calls, were discovered to be reserved exclusively for the presence of group members, and their production "exhibited characteristics previously used to argue for intentionality in gestural communication." Firstly, production of AH and WB vocalizations were more likely to accompany the arrival of a friend or a dominant than non-friends or lower-ranking individuals. In line with studies discussed above, these findings suggest that certain alarm calls are produced tactically and discriminately toward significant individuals. Secondly, signalers who produced AH and WB calls visually monitored significant individuals in the audience and continually alternated their gaze between these individuals and the snake. Thirdly, signalers persisted in their calls until all group members were a safe distance from the threat, suggesting goal- and agent-directedness in their calls. Schell et al. conclude their paper as follows: "We interpret these patterns as evidence that chimpanzee alarm calling meets the key diagnostic features of intentional signal production. Although each of these behaviors can be explained separately as the product of less complex cognitive processes, the combined overall pattern is more consistent with the hypothesis that call production is both socially directed and goal-directed." When taken alongside the plethora of complementary studies discussed thus far, this conclusion is decidedly substantive and serves to effectively refute Tomasello's hypothesis that the AG-gestures of chimpanzees are, at a foundational level, functionally different from AG-vocalizations.

Over the past three sections, I have attempted to show that the only truly salient difference between the vocalizations and gestures of chimpanzees is that the latter possess a more open-ended plasticity in production. However, because of the logical problem, the question as to whether this production is *intentional* applies equally to both modes of communication. To see this, consider the following complementary explanation of the Liebal et al. (2004) study on "attention-monitoring" during gestural communication. From a purely behavioral perspective, two correlations have been observed. First, chimpanzees observed a correlation between the production of specific gestures and an action on the part of the experimenter that leads to a reward. Second, in order to explain the physical-reorientation behavior, it is reasonable to assume chimpanzees recognize a correlation between the facial or bodily orientation of the experimenter and specific actions that, combined with gestures, tend to bring about the desired effect. Similar behavioral rules can be applied to the analogous studies on ape vocalizations. In short, these chimpanzees did not need to attribute mental states of attention and inattention to humans in order to achieve their goal of receiving food. Judging from their behavior alone, they could have merely identified and took advantage of correlations between the actions of an experimenter, their own actions, and the appearance of desired rewards.

The fact that primatologists and philosophers refer to these gestures and vocalizations as "attention-getting" is therefore in a sense functionally accurate, but as the logical problem highlights, the mere *function* of a behavior does not allow one to distinguish its underlying and/or accompanying cognitive processes. I am suggesting that if apes *do* possess a theory of mind—and there is strong evidence in favor of this—then, when combined with the fact they can use vocalizations in ways functionally equivalent to gestures, there is no salient reason to doubt that a theory of mind is operating in these circumstances as well. By Tomasello's own criteria, great apes have been observed to use vocalizations that may justifiably be referred to as "intentional signals."

6 Conclusions

Contrary to the judgment passed not only by Tomasello but many primatologists [see Arbib et al. (2008), Hammerschmidt and Fischer (2008), and Cheney and Seyfarth (2010) for reviews] chimpanzees are capable of using vocalizations in ways considerably more advanced than the inflexible behavioral displays commonly observed in the biological world. In this chapter, I have used current research to provide what I hope to be a comprehensive, bottom-up account of the communicative potential of chimpanzee vocalizations. Beginning with evidence that chimpanzee vocal production can be both socially and asocially learned, I have offered reason to believe that, like their gestures, not all chimpanzee vocalizations are genetically pre-determined and evoked by particular stimuli and emotional states. To the contrary, chimpanzees appear to have some voluntary control over *which* vocalizations they employ, *when* they employ them, and to *whom* they are directed. Chimpanzees are therefore capable of producing vocalizations with a considerable degree of contextual flexibility. This potential is amplified tremendously when smart receivers and call combinations are considered and, further, if some of their vocalizations are utilized dyadically, i.e., with sensitivity to the attentional state of the recipient. Although there is some reason to doubt this capacity, I have shown that there is no *more* reason to doubt it for gestures than for vocalizations.

I noted above that Tomasello (2008) does not present any neurobiological evidence suggestive of his view that manual gestures are produced volitionally and vocalizations are not. Tomasello's hypothesis regarding the "gestural origins" of human communication is grounded exclusively on observations of chimpanzee behavior. In this chapter, I have gone to lengths to argue that chimpanzees use gestures and vocalizations in ways that are functionally equivalent from a behavioral point of view. I have suggested that this evidence provides a foothold for a multimodal account of the evolution of human communication. There is, however, recent neurobiological evidence that both adds support this hypothesis and, as such, further contradicts Tomasello's "gestural origins" thesis. In a recent paper entitled "Chimpanzee Vocal Signaling Points to a Multimodal Origin of Human Language," Taglialatela et al. (2011) present evidence that AG-vocalizations selectively activate the Broca's area homolog in chimpanzees. The Broca's area has long been recognized as an area of the human brain critical for vocal planning and production. Taglialatela et al. note that, prior to their study, it was known that the Broca's area homolog in chimpanzees was activated by manual gestures as well as the combination of gestures and sounds. Their new study revealed that while isolated *non*-AG-vocalizations fail to get a response from the Broca's area homolog, AG-vocalizations in the absence of gestures *do* activate area of the brain. According to Taglialatela et al. (2011), "the activity observed in the Broca's area homolog reflects the production of vocal signals by the chimpanzees, (suggesting) that this critical human language region was involved in vocal signaling in the common ancestor of both modern humans and chimpanzees."

All of this suggests that the vocalizations of great apes, both in addition to and as distinct from ape gestures, can be hypothesized to have played a significant role in the evolution of human communication. Though such a discussion would clearly far exceed the space available here, I will conclude with some brief thoughts on the subject. First, the capacity to recognize individual voices on their own and in a crowd can be a powerful communicative ability, especially when combined with theory of mind. This capacity would not only aid the evolution of speech and language, but may facilitate in-group identification and therefore the evolution of specific dialects. Second, unlike gestural communication, vocalizations have the potential to facilitate communication at long distances and, as I suggested above, aid in developing idiosyncratic group identities by distinguishing the calls of neighboring communities. This ability would also presumably be important in the facilitation of "between-group" communication as it avoids the tensions that can emerge in direct physical confrontation. Thirdly, like gestures, vocal signals have been shown to be combinable to increase message complexity. While the combination of vocal signals may not be as flexible as the "simple syntax" in the gestures of "linguistic apes" proposed by Tomasello (2008), it is notable that since we presently lack the ability to effectively "decode" primate call combinations, something approximating a "simple syntax" might in fact be the case among groups of conspecifics.

As was mentioned above, there is currently a scarcity of research on great ape vocal communication (Hammerschmidt and Fischer 2008; Schel et al. 2013a). This is especially true in comparison with studies on ape gestures. Consequently, we should not overlook the fact that the current asymmetry in the research on ape gestures and vocalizations corresponds to the general perception of their respective communicative potential and roles in the origins of human communication.

References

Arbib MA, Liebal K, Pika S (2008) Primate vocalization, gesture, and the evolution of human language. Curr Anthropol 49:1053–1076

Bates E, Camaioni L, Volterra V (1975) The acquisition of performatives prior to speech. Merrill Palmer Q 21:205–226

Berlin P (2006) Voice processing in human and non-human primates. Philos Trans R Soc B 361:2091–2107

Bodamer MD, Gardner RA (2002) How cross-fostered chimpanzees (Pan troglodytes) initiate and maintain conversations. J Comp Psychol 116:12–26

Cheney DL, Seyfarth RM (1990) How monkeys see the world. University of Chicago Press, Chicago

Cheney DL, Seyfarth RM (2005) Constraints and preadaptations in the earliest stages of language evolution. Linguist Rev 22:135–159

Cheney DL, Seyfarth RM (2007) Baboon metaphysics: the evolution of a social mind. University of Chicago Press, Chicago

Cheney DL, Seyfarth RM (2010) Production, usage, and comprehension in animal vocalizations. Brain Lang 115:92–100

Clark AP, Wrangham RW (1994) Chimpanzee arrival pant-hoots: do they signify food or status? Int J Primatol 15:185–205

Crockford C, Boesch C (2003) Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. Anim Behav 66:115–125

Crockford C, Boesch C (2005) Call combinations in wild chimpanzees. Anim Behav 142:397–421

Crockford C, Herbinger I, Vigilant L, Boesch C (2004) Wild chimpanzees produce group-specific calls: a case for vocal learning? Ethology 110:221–243

Crockford C, Wittig RM, Mundry R, Zuberbühler K (2012) Wild chimpanzees inform ignorant group members of danger. Curr Biol 22:142–146

Dennett DC (1987) The intentional stance. MIT Press, Cambridge

Fehér O, Wang H, Saar S, Mitra PP, Tchernichovshi O (2009) De novo establishment of wild-type song culture in the zebra finch. Nature 459:564–568

Flombaum JI, Santos LR (2005) Rhesus monkeys attribute perceptions to others. Curr Biol 15:447–452

Galef BG (1976) Social transmission of acquired behavior: a discussion of tradition and social learning in vertebrates. In: Rosenblatt JS, Hinde RA, Shaw E, Beer C (eds) Advances in the study of behavior, vol 6. Academic Press, New York, pp 77–100

Golinkoff RM (1993) When is communication a 'meeting of minds'? J Child Lang 20:199–207

Goodall J (1986) The chimpanzees of Gombe: patterns of behavior. Harvard University Press, Cambridge

Halloran AR (2012) The song of the ape: understanding the language of chimpanzees. St. Martin's Press, New York

Hammerschmidt K, Fischer J (2008) Constraints in primate vocal production. In: Griebel U, Oller K (eds) The evolution of communicative creativity: from fixed signals to contextual flexibility. MIT Press, Cambridge, pp 93–119

Hopkins WD, Taglialatela JP, Leavens DA (2007) Chimpanzees differentially produce novel vocalisations to capture the attention of a human. Anim Behav 73:281–286

Hostetter AB, Cantero M, Hopkins WD (2001) Differential use of vocal and gestural communication by chimpanzees (Pan troglodytes) in response to the attentional status of a human (Homo sapiens). J Comp Psychol 115:337–343

Hostetter AB, Russell JL, Freeman H, Hopkins WD (2007) Now you see me, now you don't: evidence that chimpanzees understand the role of the eyes in attention. Anim Cogn 10:55–62

Janik VW, Slater PJB (1997) Vocal learning in mammals. Adv Study Behav 26:59–99

Kojima S, Izumi A, Ceugniet M (2003) Identification of vocalizers by pant hoots, pant grunts and screams in a chimpanzee. Primates 44:225–230

Krause MA, Fouts RS (1997) Chimpanzee (Pan troglodytes) pointing: hand shapes, accuracy, and the role of eye gaze. J Comp Psychol 111:330–336

Laporte MNC, Zuberbuhler K (2010) Vocal greeting behaviour in wild chimpanzee females. Anim Behav 80:467–473

Leavens D, Hopkins WD (1998) Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures. Dev Psychol 34(5):813–822

Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD (2004) Tactical use of unimodal and bimodal communication by chimpanzees, Pan troglodytes. Anim Behav 67:467–476

Leavens DA, Russell JL, Hopkins WD (2005) Intentionality as measured in the persistence and elaboration of communication by chimpanzees (Pan troglodytes). Child Dev 76:291–306

Leavens DA, Russell JL, Hopkins WD (2010) Multimodal communication by captive chimpanzees (Pan troglodytes). Anim Cogn 13:33–40

Levréro F, Mathevon N (2013) Vocal signature in wild infant chimpanzees. Am J Primatol 75:324–332

Liebal K, Pika S, Call J, Tomasello M (2004) To move or not to move: how apes adjust to the attentional state of others. Interact Stud 5:199–219

Lurz R (2011) Mindreading animals: the debate over what animals know about other minds. MIT Press, Cambridge

Marler P (1991) Song learning behavior: the interface with neuroethology. Trends Neurosci 5:199–206

Marshall AJ, Wrangham RW, Arcadi AC (1999) Does learning affect the structure of vocalisations in chimpanzees? Anim Behav 58:825–830

Mitani JC, Hunley KL, Murdoch ME (1999) Geographic variation in the calls of wild chimpanzees: a reassessment. Am J Primatol 47:133–151

Mitani JC, Nishida T (1993) Contexts and social correlates of long-distance calling by male chimpanzees. Anim Behav 45:735–746

Pollick AS, de Waal FBM (2007) Ape gestures and language evolution. PNAS. http://www.pnas.org/content/104/19/8184.long. Accessed 24 Oct 2013

Povinelli DJ, Eddy TJ (1996) What young chimpanzees know about seeing. Monogr Soc Res Child Dev 61:56–97

Povinelli DJ, Vonk J (2006) We don't need a microscope to explore the chimpanzee's mind. In: Hurley S (ed) Rational animals. Oxford University Press, Oxford

Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 1:515–526

Russell JL, Braccini S, Buehler N, Kachin MJ, Schapiro SJ, Hopkins WD (2005) Chimpanzees (Pan troglodytes) intentional communication is not contingent upon food. Anim Cogn 8:263–272

Schel AM, Townsend SW, Machanda Z, Zuberbuhler K, Slocombe KE (2013a) Chimpanzee alarm call production meets key criteria for intentionality. PLoS ONE 8(10):e76674

Schel AM, Machanda Z, Townsend SW, Zuberbühler K, Slocombe KE (2013b) Chimpanzee food calls are directed at specific individuals. Anim Behav 86(5):955–965

Seyfarth RM, Cheney DL (2003) Signalers and receivers in animal communication. Annu Rev Psychol 54:145–173

Slocombe KE, Kaller T, Turman L, Townsend SW, Papworth S et al (2010) Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. Behav Ecol Sociobiol 64:1959–1966

Slocombe KE, Zuberbühler K (2007) Chimpanzees modify recruitment screams as a function of audience composition. Proc Natl Acad Sci 104:17228–17233

Taglialatela JP, Reamer L, Schapiro SJ, Hopkins WD (2012) Social learning of a communicative signal in captive chimpanzees. Biol Lett. http://www.ncbi.nlm.nih.gov/pubmed/22438489. Accessed 24 Oct 2013

Taglialatela JP, Russell JL, Schaeffer JA, Hopkins WD (2011) Chimpanzee vocal signaling points to a multimodal origin of human language. PLoS ONE 6(4):e18852

Theall LA, Povinelli DJ (1999) Do chimpanzees tailor their gestural signals to fit the attentional state of others? Anim Cogn 2:207–214

Tomasello M (1996) Do apes ape? In: Heyes CM, Galef BG (eds) Social learning in animals: the roots of culture. Academic Press, San Diego, pp 319–346

Tomasello M (2008) Origins of human communication. MIT Press, Cambridge

Tomasello M, Zuberbuhler K (2002) Primate vocal and gestural communication. In: Bekoff M, Allen C, Burghardt G (eds) The cognitive animal: empirical and theoretical perspectives on animal cognition. MIT Press, Cambridge

Townsend SW, Deschner T, Zuberbühler K (2008) Female chimpanzees use copulation calls flexibly to prevent social competition. PLoS ONE 3(6):e2431

Wich SA, de Vries H (2006) Male monkeys remember which group members have given alarm calls. Proc R Soc B 273:735–740

Wilson ML, Hauser MD, Wrangham RW (2001) Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? Anim Behav 61:1203–1216

Winter P, Handley P, Ploog D, Schott D (1973) Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. Anim Behav 47:230–239