Animal signals and symbolism

Ulrike Griebel and D. Kimbrough Oller

The Oxford Handbook of Human Symbolic Evolution
Edited by Nathalie Gontier, Andy Lock, and Chris Sinha

Subject: Psychology, Cognitive Psychology, Cognitive Neuroscience
Online Publication Date: Mar 2021 DOI: 10.1093/oxfordhb/9780198813781.013.2

Abstract and Keywords

Human languages are symbolic. If we accept a broadly gradualist account of evolution, forerunners of the symbolism found in human languages should be observable in our closest relatives. After intensive training by humans, animals as different as great apes, dogs, sea lions, parrots, and dolphins have been shown to be able to learn, and in some cases to use, linguistic symbols with both humans and conspecifics. However, there is an absence of convincing and widely accepted evidence for symbolism in the use by non-human animals of natural communication systems in the wild. In addressing this apparent paradox, we provide definitions of fundamental differences between human symbolism and non-human communication systems and discuss foundational capacities for symbolism in non-humans. We argue that animal signals sometimes thought to resemble symbols are more likely (as proposed by Darwin) emotional expressions. We offer arguments about the evolutionary pressures that may have led to increasingly complex communication in the hominin line.

Keywords: symbolism, evolution of communication, origin of language, semantic communication, evolution and development

The gap between human and non-human communication systems

At first glance, human language and animal communication systems do not seem to have much in common. Animals of any given species have very few communication signals (within any modality), while humans have a theoretically endless repertoire of signals in either the vocal or gestural modality. Animal signals have traditionally been thought to be "fixed" (Lorenz, 1951), meaning each serves a specific communicative function, e.g., aggression (threat), courtship (fitness advertisement), warning (alarm calling), appeasement (submission or reconciliation), and so on. In contrast, human symbolic signs (words, visuo-manual signs) can be used to serve any social function imaginable, even combinations of functions such as questioning the explanation of an instruction or requesting a request for a request. The big question is: how and why did this gap between humans and other animals evolve?

Research has demonstrated that many of the capacities once thought to be uniquely human have actually evolved gradually, and can be found in one or another basic form in other animals, often in much more domain-specific ways and/or restricted contexts than in humans such as tool use, theory of mind, mental time travel, or consciousness (de Waal, 2016; Snowdon, 2004). This suggests to some scientists that the origin of symbolism and language might have relied on species-general mechanisms of learning and development such as self-organization (Elman et al., 1996). Such a viewpoint opposes the widespread saltationist idea (Chomsky, 2005), embodying the claim that language evolution evolved in an "eye blink," without precursors or convergent features found in the animal kingdom. Nevertheless, the quest for the intermediate stages of symbolic development in animals has not been as fruitful as one might have hoped. The only solid cases for animals demonstrating symbolism so far have been made in the various animal language learning studies, where several species of birds and mammals have been taught to understand and/or use human made symbols (Griebel et al., 2016). Although there are claims of symbolic usage of signals in the wild (Marler et al., 1992; Seyfarth et al., 1980), there is not a single case to our knowledge where alternative, more conservative explanations have been ruled out. In general, the evidence seems mostly consistent with Darwin's early portrayal of animal communication as emotional expression (Darwin, 1872).

The methodological problems in comparing communication in animals and humans are exacerbated by discrepancies among scientists in how they define symbolism, language, and related technical terms (semantics, reference, and contextual versus functional flexibility). These discrepancies have resulted in persistent confusion about how to interpret animal signals that appear to be language-like. We will endeavor to clarify some of these analytic terms in our exploration of the evolution of symbolism.

Signals in animal communication

We follow Maynard Smith & Harper (2003), who define a signal as an action or feature co-evolved between senders and receivers where both benefit on average from the exchange of signals. Without benefit for both participants in interaction, they argue, signals would not be selected and could not stabilize. In contrast to a signal, a "cue" is an action or state perceived by other organisms and used as a guide for action, even though it was not evolved to communicate. For example, mosquitoes use high concentrations of carbon dioxide as a cue to the presence of a mammal in the vicinity, but the carbon dioxide in the mammal's exhalations was not evolved to communicate. A signal such as an animal alarm call, on the other hand, has been shaped by natural selection to attract the attention of conspecifics, who consequently show escape behaviors. The caller gains an advantage through kin selection or reciprocal altruism (Hamilton, 1963). A cue can be selected to evolve *into* a signal. For example, in the case of an alarm call, an animal might accidentally produce an involuntary vocal cue, e.g., a gasp, when startled at perceiving a predator nearby. Conspecifics might notice the contingency between the gasp and the presence of a predator, and on the basis of that awareness, initiate escape. If the accidental vocal cue is advantageous in this way, and if it has a genetic basis, it can be shaped by natural selection through ritualization into a specialized, stereotyped, and salient signal for communication (Lorenz, 1951; Tinbergen, 1951). In this stereotyped form (a shriek perhaps) the signal becomes easily recognizable and interpretable. Such evolved signals are portrayed in the classical ethology literature as having the properties of high contrast to other signals, conspicuousness, and unambiguousness.

Animal communication signals, in this portrayal, typically evolve as “fixed signals” (Lorenz, 1951), sometimes with a typical intensity, against which variations of intensity of the signals can be understood and interpreted functionally. Examples would be canine growls, hissing of cats, alarm calls of many animals, the sexual pheromone of the silk moth (*Bombyx mori*), and the courtship displays of many animals. Fixed signals are thought to be innate, though some emerge or mature at a point when their function becomes relevant, as for example in courtship displays. Nevertheless, fixed signals may be practiced by the young in play, who presumably learn to adjust usage and intensity to a variety of contexts, as with the aggressive growling that can also occur in play fighting.

Some courtship displays, e.g., mating songs, may have to be learned individually, driven by the need to out-compete same-sex conspecifics and/or to adjust to sexual selection criteria imposed by the other sex. Both situations can lead to flexibility in the learning and use of courtship signals. Well-known examples of these learned signals are the seasonally varying songs of humpback whales and the songs of many bird species. Fixed signals in humans, such as laughter and crying, seem to appear without learning in early infancy, but humans have no problem learning to modify the forms and functions of these signals as they mature, whereas other animals do not seem to show major ontogenetic modifications in their fixed signals.

Forms and functions in communication: How to define symbolism

As a precursor to defining symbolism we need to distinguish between *form* and *function* of signals. The *function* of a sender’s signals is analogous to the “illocutionary force” of linguistic utterances (Austin, 1962). Austin employed this term to denote the social pragmatic functions of utterances, but the term is also useful to pin down crucial differences between animal signals and the spoken symbolic word, even though Austin himself never wrote about animal communication. We have adapted the term also to apply to animal and human infant communication (Oller & Griebel, 2008a, 2014). In our usage, the illocutionary force is the social-communicative function transmitted in the communicative act, e.g., aggression, appeasement, or warning. The *form* the signal takes is the act itself, in its modality or modalities (visual, acoustic, tactile, electric, chemical). The form/function pairing can be innate or learned, simple or complex, under voluntary control or not. The “perlocutionary effect” (another of Austin’s terms adapted by us for animal communication) of a signal is the change in state and/or behavior of the receiver that occurs as a result of the signal. For example, an alarm call can trigger escape, but it can also result in listeners going into hiding, freezing, or merely looking at the caller or towards the direction of possible danger, depending on the context. Thus, an alarm call does not directly cause a single specific behavioral reaction, but rather induces a state of alarm in the receiver (Owren, Rendall, & Ryan, 2010). Similarly, an aggressive signal causes the receiver to feel *threatened*, but whether the receiver will respond with aggressive signaling, physical fight, flight, submissive display, or disinterest will depend on the situation. Thus, even though the initiating signal may have a fixed illocutionary force from the perspective of the sender, the perlocutionary effect on the receiver may vary.

There are always, in principle, distinctions to be drawn between the form, the illocutionary force, and the perlocutionary effect of a signal.

This terminology helps in drawing the distinction between illocutionary force and “meaning” (the semantic content of a communication, assuming a communication has semantic content). Illocutionary forces in animal communication in the wild are interactive events initiated by a sender within a social dyad of sender and a receiver (or several receivers at once). Such illocutionary forces do not require that the signal make *reference* to anything outside the dyad. Meaning, on the other hand, does involve reference, requiring a triad: the social dyad plus the entity (object, state, activity, etc.) that the signal refers to. In triadic communication there is more than “me and you;” there is “me and you *and* that thing over there,” or “me and you *and* the idea we are talking about.” So, if a human says the word “mouse” and points to a mouse, he not only draws attention to the mouse but also invokes a name for a class of animals. Thus, when one points to a mouse and names it, one produces both an illocutionary force (naming) *and* a meaning. We can use the name/label mouse with many additional and different illocutionary functions: as aggressive criticism (“You pathetic mouse!”), as a question (“Is this a mouse?”), as an example (which we are doing in writing this), as a statement (“This is a mouse.”), as a warning (“Watch out, a mouse! Jump up on a chair … ”), as sweet talk (“You cute little mouse!”(a typical Austrian endearment)), and so on. The word “mouse” does not change its semantic content across these performances—it always invokes a category of mammal, because the word is a *symbol* for that class of entities. Symbols, the hallmark of human languages, afford unparalleled *functional flexibility* to the communication of their speakers, and every symbol has such flexibility.

Limited functions and lack of symbolism in animal communication in the wild

A fixed signal in animal communication, by contrast, is not a symbol and does not enable triadic reference nor bear semantic content. A fixed signal is thought in traditional ethology to serve a single function (e.g., aggression, or courtship), and that function cannot be reference. This viewpoint on animal signals stands in sharp contrast to the assertion in much animal literature that there exist predator-specific alarm calls, which have been portrayed as having semantic content (Seyfarth et al., 1980), or at least as permitting "functional reference" (see articles in Stegmann (2013) and commentary in Wheeler and Fischer (2012)).

The view has most famously been advocated with regard to vervet monkey alarm calls, which have been claimed to show three predator-specific types for leopards, snakes, and eagles (Seyfarth et al., 1980). In the ongoing quest for animal referential signals this was the one that inspired highest hopes—finally a candidate precursor for truly referential signaling in animals seemed to have presented itself. If one reads carefully, however, it was acknowledged in the earliest publications that the presumed alarm calls were also used in non-alarm inter—and intra-species aggressive interactions. Thus, the description "predator-specific alarm calls" was misleading because usages of the calls sometimes had nothing to do with any predator, much less a specific one. Further, a more recent study (Price, 2013) with an expanded dataset and new analyses concluded there were no sharp distinctions *among* the vervet alarm calls. All three calls seemed to be graded among each other for usage across expression of aggression or alarm, presumably being influenced by different states of arousal and thus varying in intensity of signaling. A symbol, in contrast, has a definitional requirement that it have a conventionalized and stable relation with a particular concept (at least in literal usage) and that it refer to members of the class conceptualized. Thus, the word "leopard" in English is used in referring to leopards. But descriptions of the vervet calls indicate that they do not have such stable referential content. Their use in intra-specific aggression makes clear that they *cannot refer stably to any class of predator*, and the gradation *among* the presumed calls makes clear that even in alarm circumstances, none of them can be specifically associated with just one of the presumed three predators. Others have also argued that there are specific requirements of reference/symbolization that have not been shown to occur in animal communication systems in the wild (e.g., Sinha, 2004).

A more conservative view is that alarm calls can best be understood in light of Darwin's portrayal of emotional signals (Darwin, 1872). Alarm calls can be viewed as fear/aggression expressions that encode *no* referentiality in and of themselves. In this view, an alarm call does not say or transmit anything semantically, but instead constitutes an immediate expression of an emotional state or states. These states can occur in varying degrees of arousal, with accompanying variations in intensity of subglottal pressure and consequent regime shifts in acoustic properties (especially, sudden, dramatic changes in spectral properties at threshold points of change in subglottal pressure) that may yield category-like shifts in the sounds perceived (Buder et al., 2008; Davila Ross et al., 2010). Quite a number of apparent different types of sounds can thus be produced, even if there are only a few functional/emotional dimensions along which they vary (Winter et al., 1966). The wide discrepancy in reported counts (from four to thirty-seven) of vocal types (often claimed to represent many different "meanings") in various primates (Sutton, 1979) suggests a very few functional/emotional dimensions (as few as four or five for an individual species of primate), with regime shifts accounting for perceived shifts in sound qualities along each of those dimensions.

Candidates for emotion-driven illocutionary functions of animal vocal signal systems appear to be limited to a rather short list:

1) distress expression or complaint,
2) appeasement or submission (such signals are sometimes also used in greeting),
3) greeting, contact calling, or affiliation (submission, appeasement, and/or positive excitement signals are also sometimes used for these functions),
4) positive excitement, exultation (such signals can also be used in feeding announcement),
5) aggression or threat (occurring in both intra- and interspecies interactions),
6) warning or alarm (distress or aggressive signals are often also used for this function), and
7) courtship or sexual solicitation.

With gradations of intensity and accompanying regime shifts in acoustics these functions can be expressed with quite a diversity of perceived sounds. Furthermore, the emotional states may mix—an animal may experience more than one emotion simultaneously, for example, resulting in gradations and regime shifts both *within and across* calls, yielding varying degrees of, for example, aggression and alarm, a possibility suggested by the gradations described in Price (2013) for vervet monkey calls. If this conservative view is correct, semantics is simply not implicated in these signals, only illocutionary forces driven by different flavors of emotion, varying in intensity, and mixing.

In humans, conflicting emotions are often revealed in facial expressions that are controlled and recombined by forty-two different muscles (Eckman, 1994). Gradations and conflicting emotions can either be expressed in bilaterally asymmetrical facial displays or differences/contradictions of upper and lower facial displays, e.g., the surprise face shows the wide-open eyes of the fear face, but not the clenched teeth of the fear face. Instead surprise is accompanied usually by a playful open mouth of a friendly face. And indeed, surprises can go either way, pleasant or unpleasant! Human languages have ways to express conflicting emotions either through semantic content (with flat intonation, “he’s a bloody saint”) or prosody (with sarcastic emphasis on the last word, “this is just great!”).

Tinbergen (1952) described how animals with conflicting emotions sometimes display displacement behaviors: e.g., two skylarks engaged in furious combat suddenly both peck at the ground as if they are feeding. Similarly, during mating rituals, just before the actual mating, some birds start to preen themselves, presumably because the motivation to mate is conflicted. Humans often show behaviors like pacing, grooming (e.g., scratching, twisting rings, chewing nails) or surrogate feeding (chewing a pen, smoking) in situations of conflicting emotions.

Many vocalizations in animals and humans may also involve mixed emotions. This helps explain discrepancies among reports on vocal repertoires, specifically in social animals where nuances of vocal emotional expression must be under selection pressure to facilitate social living. Alarm calling in monkeys and other animals appears to be so complex because of gradations and mixing of emotions, and as a result there may be no easy mapping of signal to function, and listeners (either human observers or conspecifics) may need to apply intelligence to the task of interpretation.

The notion of perlocutionary effect helps here, because it highlights the role of the animal receiver in interpreting fear/aggression signals intelligently. The receiver can be induced, in this view, to make an intelligent decision based on a sender's alarm signal, not because the sender names a particular predator (it does not), but because the signal induces emotions and occurs in a circumstance that can help guide interpretation by the receiver. Receivers may have the capacity to learn the kinds of call properties that are associated with likely predator types and may indeed determine that a particular mode of escape is warranted based on that knowledge. The receiver's perlocutionary response thus seems very likely to be heavily influenced by fear induction, the natural result of hearing a vocalization of a fearful conspecific, followed by action informed by prior experience. This more conservative stance, rejecting referentiality (or at least withholding judgment about it) with regard to animal alarm calls in the wild, has been advocated in a variety of publications where other explanations of the proposed interpretations of alarm calls by animals can be found (Rendall et al., 2009; Owings & Morton, 1998; Owren & Rendall, 2001; Owren et al., 2010; Oller & Griebel, 2014).

Even if we were to suppose that there *are* three categorically distinct vervet alarm calls each used exclusively to warn about one of three different predators, the calls would still not qualify as symbolic if they are limited to expressing warning. Symbolic signals can be used to express *any possible* illocutionary function, as indicated above. If their signals were symbolic, vervets would be able to use the "leopard call" with an intention *other than* warning, for example, as an unemotional statement about a leopard, a question about a leopard, or an invocation of the leopard seen last week. A related matter is that animal signals have not been proven to reverse from positive to negative valence on different occasions nor to serve vastly different purposes. So, an animal warning or alarm signal has never been reported and presumably cannot be used to signal appeasement or social affiliation. An additional critical limitation of function is that animal signals in the wild never, as far as we can tell, simply refer to or comment on something. In contrast, a typical human eighteen-month old often names objects, foods, or body parts with no intention other than naming. Every linguistic symbol has this property—it can be used just for naming (labeling). Animal signals do not name, but instead show a mapping of a signal to an emotionally driven function or mixture of such functions.

Suppose at some point we find animals that can be proven to have categorically distinctive alarm calls for different predators (vervets do not appear to qualify). We would still have to show that each categorical signal could be used to serve multiple illocutionary functions, including the function of naming for no purpose other than naming; only then could we argue convincingly for semanticity of the signals. Thus the possible diversification of the relation between signal and functions in animal communication remains doubtful and at least disputable. To our knowledge, essentially the same concerns expressed here about claims of symbolism and semanticity in vervets apply to all the species that have been claimed to possess predator-specific calls (see e.g., Arnold & Zuberbühler, 2006; Gozoules et al., 1995; Macedonia & Evans, 1993; Suzuki, 2012; Griesser, 2009; Manser et al., 2002; Furrer & Manser, 2009; Hollén & Manser, 2007).

Decoupling of form and function in human communication

Human infants show decoupling of signal and function very early in development, even in their pre-speech sounds (protophones) during the very first months of life (Jhang & Oller, 2017; Oller et al., 2013). Human infants produce a variety of protophones from birth (Dominguez et al., 2016; Nathani et al., 2006), and recent evidence suggests even preterm infants, at thirty-two weeks gestational age, still in neonatal intensive care, produce protophones (Caskey & Vohr, 2013, Oller et al., 2014). By three months of age these include squeals, growls, vowel-like sounds, and raspberries, for example, and all of these are produced with a full range of valence, showing positive, neutral, and negative facial affect. “Crying.” on the other hand, which appears to be a fixed signal of distress (caused by pain or hunger), occurs from birth, but is bound to negative affect, even though the range of negative expressions grows as the infant matures (Green et al., 2011). In adulthood, humans show even more remarkable vocal flexibility, and although crying continues to reveal strong emotion, it can express relief and can occur in joyful situations.

Functional flexibility in human languages is not restricted to the vocal domain; sign languages show similar flexibility (Lyons, 1991; Stokoe, 1960). Interestingly, human body language and other natural or iconic gestures are not entirely stereotyped to serve fixed functions either. For example, a normally aggressive iconic hand gesture suggesting the shooting of a pistol can be used to transmit a variety of illocutionary forces, as for example a serious threat (you are dead), a part of a story narrative (then he did this …), self-deprecation (how silly of me!), and so on. Subconscious human body language apparently conveys information mostly encoded in timing and rhythm, rather than through stereotyped or consciously learned movements (Grammer, 1995; Grammer et al., 1999; Grammer et al., 2000). Substantial research on gestural communication of great apes in the wild suggests considerable gestural activity and learning, but thus far there is no evidence of semanticity in their natural communication systems (Call & Tomasello, 2007). On the other hand, human trained animals of many taxa clearly show semantic learning assuming their training begins early in life and is maintained consistently (see review in Griebel et al. (2016)).

Distinguishing function and context

Another important issue causing confusion in animal communication research concerns the distinction between *context* and *function*. Since animals don’t talk, we have to infer functions of signals (e.g., growl) from context (e.g., physical fight). But a growl may be used in a variety of contexts, e.g., during feeding (competition for food), traveling (leadership disputes), or courtship (competition for mates), and it seems likely that such growls have a single illocutionary function of aggression. Similarly, even in a single context, such as feeding, different vocal signals with different functions can occur—for example, a social affiliation signal or an aggressive one. Thus, it is not possible to determine function by context alone. Still, published research has argued for functional flexibility in ape vocalizations on the basis of contextual flexibility alone (Clay et al., 2015). Furthermore, since simultaneous experience of multiple emotions appears to be possible, we cannot always determine what emotion or emotions are involved in motivating a particular vocalization.

We discussed the usage of aggressive and/or distress vocalizations for alarm purposes earlier. The term "functional reference" has been proposed to be a more conservative and accurate description of differentiated animal alarm calls, presumably in an effort to preserve the claim that such calls suggest foundations for semantics, if not semantics per se. Playback experiments that show tendencies for animals in various species to implement escape behaviors that correspond at better than chance levels to the predators presumed to be associated with the calls played back are touted as empirical evidence supporting the notion of functional reference. However, by the same kind of reasoning, we would have to conclude that the human infant cry is functionally referential, since mothers determine that cries are indicators of, for example, hunger, pain, or being wet, on the basis of differences in the acoustics of cries, along with intelligent assessments of the infant's current state (see the full argument in Oller & Griebel, (2014)). The human infant cry is, however, patently non-semantic. Like animal signals in the wild, as far as is known, human infant cry is an emotional expression that can be interpreted in a variety of ways by intelligent listeners, whose perlocutionary reactions can be very systematic even though the vocalizations themselves do not encode semantic/symbolic content.

If animals in the wild do not use symbols, where do we find the foundations of human symbolism?

While animals in the wild do not appear to employ symbols, with human training, language-like symbolic communication can indeed be learned by a variety of animals. Human-trained animals of many taxa clearly show semantic learning if their training begins early in life and is maintained consistently (Griebel et al., 2016; Lyn, this volume). As in the case of human infants, symbol usage (which in trained animals often takes the form of manual-visual signs or visual symbols on a communication board) does not begin immediately, however. This fact hints at the likelihood that animals go through similar steps in learning as those required for human infants (e.g., signal exploration and play, functionally flexible expression with the signals, face-to-face turn-taking interaction using the signals, joint attention, imitation of signals, and perhaps others) *before* symbolic learning of words can occur (Oller, 2000). But animals *do* learn to use symbols, so the foundations for language-like symbolization are clearly present in at least some groups of animals.

A necessary foundation for symbolism, the categorization of objects and other entities, is clearly present in animals. A vast array of studies shows that animals recognize groupings of objects and other entities, respond to them systematically, and can learn new ones (Griffin, 1992; de Waal, 2016). The capacity for categorization is crucial because, of course, it is categories that symbols denote.

Another important basis for symbolism is having signals that are learned, rather than innate, and that show at least some diversity of function (Griebel & Oller, 2008). For example, mating displays advertise health and strength corresponding to the signaler's genes and thus function as fitness indicators. We find elaborate vocal mating displays in many songbirds and marine mammals such as the humpback whale, where songs are learned from conspecifics. There is sometimes a premium on accuracy of learning a mating song, and in other cases songs are modified creatively over time, as in humpback whales, where the song is changed with every mating season to impress females (Payne & Payne, 1997; Payne & McVay, 1971). These songs often have minimal units, comparable to syllables or notes (Marler & Slabbekorn, 2004), which are recombined and repeated in systematic ways. Some species even show an increasing repertoire over the years, while others seem to have a stable repertoire throughout their lifespan (Baptista & Petrinovich, 1986; Nottebohm, 1981; Payne & Payne, 1997; West et al., 1997). Such songs suggest a high degree of signal flexibility in some species.

In seeking to determine functions of singing in animals, it is often difficult to discriminate between courtship and territorial songs within species, since both are fitness displays and use the same vocal material. Natural selection of songs is based on perlocutionary effects in females who are attracted to, and male competitors who are repelled from, the singer's territory. We might view such songs as having either one social function, fitness advertisement, or two different ones, courtship and territorial defense. In either case, animal songs are not used for other social functions such as warning, appeasement, or greeting.

Another possibility for flexible, diversified functions of songs concerns a kind of deception, proposed by Dawkins and Krebs (1978) as the "Beau Geste" hypothesis, intended to explain why some songbirds copy songs of other species. A male mockingbird, for example, can make it seem that a territory is full of birds by producing diverse songs from various species, thus discouraging competing males from entering the territory. It has also been claimed that some birds imitate predator calls to discourage food or mate competition (Greene & Meagher, 1998).

Deception has also been reported to occur in the use of non-mating calls and other actions (Byrne & Whiten, 1990; Gibson, 1990; Menzel, 1988; Mitchell & Thompson, 1986; Savage-Rumbaugh & MacDonald, 1988). For example, an individual might produce an alarm call, seemingly to make conspecifics seek cover, leaving a newly found food cache to himself. Deceptive signaling is hard to prove in animal research, not only because of multiple possible interpretations of an animal's motivations (perhaps the individual actually thought he perceived a predator, mistakenly), but because it necessarily has to occur rarely, or else it loses its effectiveness.

Thus, although deception is hard to prove, it seems likely to happen in animals at low frequency. One line of reasoning about the likelihood that intentional deception occurs in animals relies on the fact that humans use deception, and it seems unlikely that it has evolved only once or without precursors. In any case, deception can be seen as an occasional diversification of the function of a signal, even if the function of the signal is interpreted erroneously by the receiver. Deception is a special case of communication that can be viewed as parasitic because it depends upon an existing signaling system with stable relations between signals and functions; it is this otherwise stable system that is exploited for deception, and rarity of deception is thus required. To the extent that intentional deception occurs in animals, it suggests a foundation for flexible signaling.

Another circumstance where evolution has produced varied and more flexible signal repertoires involves sociality. Social living demands extensive communication among individuals to minimize conflict, coordinate activities, to bind individuals together (“social cohesion”), and to establish boundaries between different social groups as in the case of calling “dialects.” We find examples of social cohesion calls in cetaceans (whales and dolphins, see Kuczaj and Macheka (2008)), elephants, primates, some social bats, and birds (Snowdon, 2004).

For example, orcas (killer whales) live in stable groups called pods and produce a variety of vocal signals other than echolocation clicks. Tonal whistles and pulsed calls can be either very discrete or highly flexible (Ford, 1991). Each pod has a group-specific repertoire of discrete calls that every member of the pod is able to reproduce and that is stable over long periods. Related groups share some of the same calls, but overlap is not complete. Each call type occurs in various contexts (feeding, resting, traveling, etc.). It is thought that the calls indicate different emotional states and moods, and that they function in coordination of activities and in locating members of the group. So far there has been no consistent relationship demonstrated between signal type and emotional states (e.g., aggressive signals, fear signals, mating signals, etc.) in orcas.

Dolphins seem to have vocalization repertoires similar to those of orcas, even though their social systems are more like the fission-fusion type found in chimpanzees (McCowan & Reiss, 1995; McCowan & Reiss, 1997). To date, only a few species of dolphins have been studied sufficiently to know. For bottlenose dolphins, a unique type of vocal signal has been claimed to exist, the “signature whistle.” It is said to be learned by each animal over the course of a few months to a few years, being unique to each individual (Caldwell & Caldwell, 1965; Tyack, 2000; Weiβ et al., 2006; Sayigh et al., 2007). However, these signature whistles are also produced by other individuals in the group, seemingly contradicting their very definition. It is not clear, then, that the proposed signature whistles are not merely group repertoire signals of the sort found in orcas (McCowan & Reiss, 2001). Advocates of the signature whistle idea argue in response to this objection that group members imitate each other’s signature whistles, but experiments that would clearly distinguish between the family repertoire hypothesis and the signature whistle hypothesis have not yet been conducted. The proposed advantage of signature whistles is that dolphins recognize group members by their signature whistles while hunting in murky waters or coordinating other activities under limited visibility. It is argued that distinctiveness of voice as a cue to an individual's identity is likely to be unreliable under water, partly because of compression of vocal structures during diving (Tyack & Clark, 2000). Nevertheless, claims for signature whistles have also been made for parrots who vocalize in air without the above mentioned restrictions (Berg et al., 2012), but again, it has not been demonstrated that a "family repertoire" hypothesis can be excluded. At present the signature whistle hypothesis is an interesting concept, but it needs to be confirmed with decisive experimental work. Demonstration of truly symbolic communication must include showing that a particular signal can be used with different illocutions (perhaps displaying different affect) on different occasions. For example, a particular signal could be shown to be a symbol if it were used on separate occasions: 1) as an indicator that the sender needed help, 2) merely to indicate her location, 3) to invoke an individual that is absent at the moment, etc. The demonstration would only be effective if the different illocutions could be shown to be followed by different behavioral responses of receivers (perlocutions). For example, the receiver might: 1) approach and help the sender, 2) might look in the direction of the sender, perhaps calling in return to indicate her own location, or 3) initiate search behaviors for the animal invoked by the sender's vocalization, etc. We would not be surprised if dolphins are capable of all these things, but empirical demonstration is required.

There have been reports of fixed signals with specific functions in dolphins as well, for example, for aggression and courtship (Connor & Smolker, 1996; Herzing, 1996; McCowan & Reiss, 1995; Kuczaj & Eskelinen, 2014). Blomqvist et al. (2005) even suggest a signal equivalent to human laughter in dolphins, which is used in the context of play. But in general, it has been challenging to ascribe fixed signal character to cetacean vocalizations because of their variability, which is associated with the fact that learning is extensively involved. In fact, all aspects of signaling in cetaceans have been difficult to study simply because it is hard to track and record animals in water and to determine which individuals in a group are vocalizing.

From a certain perspective, camouflage can be viewed as the opposite of signaling: an organism evolves camouflage to prevent cueing in order to avoid detection by predator or prey. Although it does not evolve to benefit the sender as well as the receiver, co-evolution between sender and receiver nevertheless can occur, with a resulting arms race of camouflage and detection. Camouflage acquires interest in the evolution of flexible signaling due to the remarkable case of camouflage in cephalopods (octopus, squid, nautilus), animals that produce extremely versatile and fast changing color patterns on their skin, and where camouflage patterns are produced by the same basic chromatophore elements as their social communication signal patterns. It has been speculated that the flexible patterns evolved first for camouflage. In support of this idea, we find these flexible color patterns in social and solitary species alike, e.g., common octopus, which is solitary, and the Caribbean reef squid, which is social. The color patterns seem to be innate, but mature during development, showing gradation and stereotypy very much like acoustic communication signals. The skin patterns are used for aggression and courtship as well as for camouflage (Hanlon & Messenger, 1996; Messenger, 2001, Moynihan & Rodaniche, 1982; Byrne et al., 2003).

The emergence of symbolization in humans

There are, then, numerous indications that animals of extremely diverse taxa have capacities that form foundations for various aspects of symbolism: categorization, flexible signal production, and learning. In the following, we present speculations about how the human line diverged from other animals (especially other primates) in evolving more extensive foundations, step by step, for a symbolic language capacity (see Arbib, this volume).

If humans had taken a primarily gestural route to language, our task here would seemingly be easier, because the great apes have considerable voluntary hand control and can learn new gestures fairly easily (Fouts, 1987; Gardner et al., 1989; Premack, 1971). In fact, it has been argued that a gestural route was the initial one, and that after some considerable evolution of language-like capabilities in gesture, the vocal route took over in the hominin line (Hewes, 1992; Tomasello, 2008). We are skeptical of this hypothesis, in part because we note that no primate group has evolved symbolic communication in the wild, even in the gestural domain, and in part because we take the pattern of development of very early communication in the human infant as indicative of the evolutionary stages of language.

Within in the first six months of human life, communication is extremely prominent through both protophone vocalization and facial affect, while communicative signals with hand gestures or movements of the head are essentially absent (Burkhardt-Reed et al., 2017). Throughout the first year of life vocalizations are used communicatively, but gesture becomes communicative (especially in the form of pointing) only in the second half year. Further, every human social group with normal hearing uses spoken language as its primary form of communication. Thus, whatever the early role of gesture might have been, and in spite of its secondary role in human communication (especially symbolic communication) currently (Iverson & Goldin-Meadow, 1998), vocalization is clearly the predominant mode. The other great apes are extremely limited in the ability to learn vocalization or to use it flexibly (Ackermann et al., 2014;, with minor exceptions (Hopkins et al., 2011; Hopkins et al., 2007; Clay et al., 2015; Crockford & Boesch, 2003), suggesting that hominins must long ago have come to possess a capability and an inclination to produce vocalization copiously and with notable flexibility. Since such a capability is observable from the first months in human infants (Jhang & Oller, 2017; Oller et al., 2013), and all other features of vocal language depend upon it (Oller et al., 2016), we reason that ancient hominins must have come under selection pressure to produce voluntary and functionally flexible vocalizations not long after their divergence from other apes.

One important factor that must have influenced early hominin evolution, which is actually a foundational capacity of all flexible vocal signaling, is the evolution of more extensive voluntary breath control, a necessary foundation for voluntary control of the glottis, and thus for voluntary phonation. Inquiring into how humans might have evolved glottal control, we turn to comparative evidence. In non-humans, vocal flexibility (especially in the form of learning of new vocal types) is found notably in marine mammals (Gisiner & Schusterman, 1992; Schusterman et al., 1992; Tyack & Sayigh, 1997). Living at least intermittently in water and feeding from water, voluntary breath control is required for diving and foraging. It is conspicuous that overwhelmingly, the mammals that have been reported to show notable vocal learning, and these reports include elephants (Stoeger & Manger, 2014), have aquatic histories. Consistent findings of seafood and fish in hominin sites point toward waterside living in much of hominin history (see review in Verhaegen, et al., (2007). The hypothesis that ancient hominins evolved under the influence of waterside living has been on the table for decades (Hardy, 1960), and considerable evidence has been amassed in support of it (see e.g., Kuliukas, 2011). Initial reception was generally negative within paleoanthropology (Langdon, 1997), although some paleoanthropologists have more recently expressed support for waterside influences on special human features such as bipedalism (Tobias, 2011; Wrangham et al., 2009). The waterside hypothesis has had much more positive reception in other academic domains (Dennett, 1995). The idea is not going away because it offers an explanation for a wide range of differences between humans and other primates, among others, relative hairlessness, large amounts of subcutaneous fat, and voluntary breath control, all of which are very common in amphibic and marine mammals. Even bipedalism in hominins acquires a straightforward explanation under the waterside scenario, because wading during foraging clearly fosters upright gait (Kuliukas et al., 2009). Evidence has been supplied for a great many physical and behavioral features of humans that are consistent with evolution under the influence of waterside living (Morgan, 1997; Verhaegen et al., 2011; Niemitz, 2010; Schagatay, 2011; Joordens et al., 2019; Wood, 2019).

For present purposes the main issue is that humans must at some point have acquired an extremely flexible capacity for voluntary breath control, and this capacity could have put ancient hominins in an especially responsive position with respect to selection pressure on voluntary vocalization. Developmental evidence becomes relevant here. Human infants are relatively altricial (born relatively prematurely) among mammals and very altricial among primates (see Suman, this volume). They are born with brains that are relatively small compared to their ultimate size, and slow physical development results in human infants not being able to forage on their own until age seven to eight. Thus human infants need parental investment for twice as long as the infants of other apes (Locke & Bogin, 2006). These facts suggest selection pressure in hominin infants, who would have been increasingly altricial across the evolutionary timeframe, to signal physical fitness to their caretakers. It has been proposed that such fitness signaling, while it could have been focused in a variety of modalities, was especially evolved in copious flexible vocalizations (Locke, 2006; Oller & Griebel, 2008b). The flexible vocalization capacity, beginning as a fitness signal, would have created an environment where caregivers would have come under selection pressure to evaluate the flexibility of the signals, and would thus have been under selection pressure to interact in turn-taking with their infants. With selection on vocal capacity, both infants and mature individuals would have been in an increasing position to utilize vocalization socially, and with increasing group sizes in ancient hominins

(Dunbar, 1993, 1996), pressures for silence (to avoid alerting predators) would have been counteracted. Alliance formation, mating solicitation, and all other forms of social interaction could have come to involve vocalization to a much greater extent than in any other ape. The capability for flexible vocalization in modern humans is based on refined monosynaptic projections of motor cortex to the brainstem nuclei controlling laryngeal muscles (Ackermann et al., 2014; Jürgens, 1992). Thus far, such foundations for vocalization have not been found in other primates. It seems likely that the evolution of the brain mechanisms allowing flexible vocalization were under selection pressure for most of hominin history.

In such a scenario, flexible signals in the form of a wide variety of vocalizations already being utilized for social purposes in interaction would have come to represent raw material for the evolution of symbolization. Community living, tool use, group hunting and foraging, as well as food sharing, all enhanced by vocal communication in illocutionary form, could have provided a basis for triadic interactions involving joint attention. While we think that flexible vocalizations in humans could have evolved as early several million years ago, it is difficult to gauge when true symbols might have emerged in humans. In any case, the emergence of symbols would have required the important steps outlined earlier as precursors.

References

Ackermann, H., Hage, S. R., & Zeigler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences 37*, 529–546.

Arnold, K., & Zuberbühler, K. (2006)). The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour 72*, 643–653.

Austin, J. L. (1962). *How to do things with words.* London: Oxford University Press.

Baptista, L. F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow: Social factors and sex differences. *Animal Behavior 34*, 1359–1371.

Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R. & Bradbury, J. W. (2012). Vertical transmission of learned signatures in a wild parrot. *Proc. R. Soc. B. doi:10.1098/rspb.2011.0932*.

Blomqvist, C., Mello, I., & Amundin, M. (2005). An acoustic play-fight signal in bottlenose dolphins (*Tursiops truncatus*) in human care. *Aquatic Mammals 31*, 187–194.

Buder, E. H., Chorna, L., Oller, D. K., & Robinson, R. (2008). Vibratory regime classification of infant phonation. *Journal of Voice 22*, 553–564.

Burkhardt-Reed, M., Griebel, U., & Oller, D. K. (2017). *The relative roles of voice, face, and gesture in early communication development: Predominance of face and voice in the* *first half-year of communicative development*. Los Angeles, CA: American Speech-Language-Hearing Association.

Byrne, R. A., Griebel, U., Wood, J. B., & Mather, J. A. (2003). Squid say it with skin: A graphical model for skin displays in Caribbean Reef Squid (*Sepioteuthis sepioidea*). In K. Warnke, H. Keupp., & S. V. Boletzky (Eds.), Coleoid cephalopods through time. *Berliner Paläobiologische Abhandlung 3*, 29–35.

Byrne, R. & Whiten, A. 1990. Tactical deception in primates: The 1990 database. *Primate Report 27*, 1–101.

Caldwell, M. C., & Caldwell, D. K. (1965). Individualized whistle contours in bottlenose dolphins (*Tursiops truncatus*). *Science 207*, 434–435.

Call, J., & Tomasello, M. (2007). *The gestural communication of apes and monkeys.* New York: Taylor & Francis Group/Lawrence Erlbaum Associates.

Caskey, M., & Vohr, B. R. (2013). Assessing language and language environment of high risk infants and children: A new approach. *Acta Pediatrica 102*, 451–461.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry 36*, 1–22.

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ 3*, 1–20. **doi.org/10.7717/peerj.1124**

Connor, R. C., & Smolker, R. A. (1996. “Pop” goes the dolphin: A vocalization male bottlenose dolphins produce during consortships. *Behaviour 133*, 643–662.

Crockford, C., & Boesch, C. (2003). Context-specific calls in wild chimpanzees, Pan troglodytes versus: Analysis of barks. *Animal Behaviour 66*, 115–125.

Darwin, C. (1872). *The expression of emotions in man and animals.* Chicago: University of Chicago Press.

Davila Ross, M., Owren, M. J., & Zimmermann, E. (2010)). The evolution of laughter in great apes and humans. *Communicative & Integrative Biology 3*, 191–194.

Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation? In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (pp. 382–409). Sunderland: Sinauer Associates.

de Waal, F. (2016). *Are we smart enough to know how smart animals are?* New York: Norton.

Dennett, D. (1995). *Darwin’s dangerous idea.* London: Penguin.

Dominguez, S., Devouche, E., Apter, G., & Gratier, M. (2016). The roots of turn-taking in the neonatal period. *Infant and Child Development 25*(3), 240–255.

Dunbar, R. (1993). Coevolution of neocortical size, group size, and language in humans. *Behavioral and Brain Sciences 16*, 681–735.

Dunbar, R. (1996). *Gossiping, grooming and the evolution of language*. Cambridge, MA: Harvard University Press.

Eckman, P. (1994). All emotions are basic. In P. Eckman & R. J. Davidson (Eds.), *The nature of emotion.* New York: Oxford University Press.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Ford, J.K.B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. Canadian Journal of Zoology 69: 1454–11483.

Fouts, R. S. (1987). Chimpanzee signing and emergent levels. In G. Greenberg & E. Tobach (Eds.), *Cognition, language and consciousness: Integrative levels.* (pp. 57–84) Hillsdale, NJ: Lawrence Erlbaum Associates.

Furrer, R. D. & Manser, M. B., (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *The American Naturalist 173*(3), 400–410.

Gardner, R. A., Gardner, B. T., & Van Cantfort, T. E. (Eds.) (1989). *Teaching sign language to chimpanzees.* Albany, NY: Suny Press.

Gibson, K. R. (1990). Tool use, imitation, and deception in captive cebus monkey. In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives.* (pp. 205–218). New York: Cambridge University Press.

Gisiner, R., & Schusterman, R. J. (1992). Sequence, syntax, and semantics: Responses of a language-trained sea lion (*Zalophus californianus*) to novel sign combinations. *Journal of Comparative Psychology 106*, 78–91.

Gozoules, H., Gozoules, S., & Ashley, J. (1995). Representational signaling in non-human primate vocal communication. In E. Zimmerman, J. D. Newman, & U. Jürgens (Eds.), *Current topics in primate vocal communication.* (pp. 235–252). New York: Plenum Press.

Grammer, K. (1995). *Signale der Liebe. 3. neu überarbeitete Auflage,* Munich: dtv-Wissenschaft.

Grammer, K., Honda, R., Schmitt, A., & Juette, A. (1999). Fuzziness of nonverbal courtship communication: Unblurred by motion energy detection. *Journal of Personality and Social Psychology 77*, 509–524.

Grammer, K., Kruck, K., Juette, A., & Fink, B. (2000). Non-verbal behavior as courtship signals: The role of control and choice in selecting partners. *Evolution and Human Behavior 21*, 371–390.

Green, J. A., Whitney, P. G., & Potegal, M. (2011). Screaming, yelling, whining and crying: Categorical and intensity differences in vocal expressions of anger and sadness in children's tantrums. *Emotion 11*, 1124–1133.

Greene, E., & Meagher, T. (1998). Red squirrels, *Tamiasciurus hudsonicus*, produce predator-class specific alarm calls. *Animal Behavior 55*(3), 511–518. **doi.org/10.1006/anbe.1997.0620**

Griebel, U. & Oller, D. K. (2008). Evolutionary forces favoring contextual flexibility. In D. K. Oller & U. Griebel (Eds.), *Evolution of communicative flexibility: Complexity, creativity and adaptability in human and animal communication*.(pp. 9–40). Cambridge, MA: MIT Press.

Griebel, U., Pepperberg, I. M., & Oller, D. K. (2016). Developmental plasticity and language: A comparative perspective. *Topics in Cognitive Science 8*, 435–445.

Griesser, M. (2009). Mobbing calls signal predator category in a kin group-living bird species. *Proceedings of the Royal Society B 276*, 2887–2892.

Griffin, D. R. (1992). *Animal minds.* Chicago: University of Chicago Press.

Hamilton, W. D. (1963. The evolution of altruistic behavior. *American Naturalist 97*, 354–356.

Hanlon, R. T., & Messenger, J. B. (1996). *Cephalopod behaviour,* Cambridge: Cambridge University Press.

Hardy, A. (1960). Was man more aquatic in the past? *New Scientist 7*, 642–645.

Herzing, D. L. (1996). Vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*. *Aquatic Mammal 22*, 61–79.

Hewes, G. W. (1992). Primate communication and the gestural origin of language. *Current Anthropology 33*, 65–84.

Hollén, l. I., & Manser, M. B. (2007). Motivation before meaning: Motivational information encoded in meerkat alarm calls develops earlier than referential information. *The American Naturalist* 169(6), 758–767.

Hopkins, W. D., Taglialatela, J. P., & Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour 73*, 281–286.

Hopkins, W. D., Taglialatela, J., & Leavens, D. A. (2011). Do apes have voluntary control of their vocalizations and facial expressions? In A. Vilain, J. L. Schwartz, C. Abry, & J. Vauclair (Eds.), *Primate communication and human language.* (pp. 71–88). Amsterdam: John Benjamins Publishing.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture as they speak. *Nature 396*, 228.

Jhang, Y., & Oller, D. K. (2017). Emergence of functional flexibility in infant vocalizations of the first 3 months. *Frontiers in Psychology 8*, Art. 300. **doi.org/10.3389/fpsyg.2017.00300**

Joordens, J. C. A., Feibel, C. S., Vonhof, H. B., Schulp, A. S., & Kroon, D. (2019). Relevance of the eastern African coastal forest for early hominin biogeography. *Journal of Human Evolution 131*, 176–202.

Jürgens, U. (1992). On the neurobiology of vocal communication. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication.* (pp. 31–42). New York: Cambridge University Press.

Kuczaj, S. A. & Eskelinen, H. C. 2014. Why do Dolphins Play? *Animal Behavior and Cognition,* 1, 113–127.

Kuczaj, S. A., & Macheka, R. (2008). The role of play in the evolution and ontogeny of contextually flexible communication. In D. K. Oller & U. Griebel (Eds.), *Evolution of communicative flexibility: Complexity, creativity, and adaptability in human and animal communication.* (pp. 253–277). Cambridge, MA: MIT Press.

Kuliukas, A. V. (2011). Langdon's critique of the aquatic ape hypothesis: It's final refutation, or just another misunderstanding? In M. Vaneechoutte (Ed.), *Was man more aquatic in the past? Fifty years after Alister Hardy—Waterside hypotheses of human evolution.* Sharjah, UAE: Bentham Science.

Kuliukas, A. V., Milne, N., & Fournier, P. (2009). The relative cost of bent-hip bent-knee walking is reduced in water. *Homo: The Journal of Comparative Human Biology 60*, 479–488.

Langdon, J. (1997). Umbrella hypotheses and parsimony in human evolution: A critique of the aquatic ape hypothesis. *Journal of Human Evolution 33*, 479–494.

Locke, J. L. (2006). Parental selection of vocal behavior: Crying, cooing, babbling, and the evolution of language. *Human Nature 17*, 155–168.

Locke, J., & Bogin, B. (2006). Language and life history: A new perspective on the evolution and development of linguistic communication. *Behavioral & Brain Sciences 29*, 259–325.

Lorenz, K. (1951). Ausdrucksbewegungen höherer Tiere. *Naturwissenschaften 38*, 113–116.

Lyons, J. (1991). *Natural language and universal grammar.* Cambridge: Cambridge University Press.

Macedonia, J. M., & Evans, C. S. (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology 93*, 177–197.

Manser, M. B., Seyfarth, R., & Cheney, D. (2002). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences 6*, 55–57.

Marler, P., Evans, C. S., & Hauser, M. (1992). Animal signals: Reference, motivation or both? In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication.* (pp. 66–86). New York: Cambridge University Press.

Marler, P., & Slabbekorn, H. (2004). *Nature's music: The science of birdsong.* Amsterdam: Elsevier Academic Press.

Maynard Smith, J., & Harper, D. (2003). *Animal signals.* Oxford: Oxford University Press.

McCowan, B., & Reiss, D. (1995). Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): A re-evaluation of the signature whistle hypothesis. *Ethology 100*, 193–209.

McCowan, B., & Reiss, D. (1997). Vocal learning in captive bottlenose dolphins: A comparison with humans and nonhuman animals. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development.* (pp. 178–207). New York: Cambridge University Press.

Mccowan, B., & Reiss, D. (2001). The fallacy of "signature whistles" in bottlenose dolphins: A comparative perspective of "signature information" in animal vocalizations. *Animal Behaviour 62*, 1151–1162.

Menzel, E. 1988. A group of young chimpanzees in a one-acre field: Leadership and communication. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence.* (pp. 155–159). Oxford: Clarendon Press.

Messenger, J. B. (2001). Cephalopod chromatophores: Neurobiology and natural history. *Biological Review 76*, 473–528.

Mitchell, R. W., & Thompson, N. S. (1986). *Deception: Perspectives on human and nonhuman deceit.* Albany, NY: State University of New York Press.

Morgan, E. (1997). *The aquatic ape hypothesis.* London: Souvenir Press.

Moynihan, M. H. & Rodaniche, A. F. (1982). The behaviour and natural history of the Caribbean reef squid *Sepioteuthis sepioidea* with a consideration of social, signal and defensive patterns for difficult and dangerous environments. *Advances in Ethology 125*, 1–150.

Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics and Phonetics 20*, 351–367.

Niemitz, C. (2010). The evolution of the upright posture and gait—a review and a new synthesis. *Naturwissenschaften 97*, 241–263.

Nottebohm, F. (1981). A brain for all seasons: Cyclical anatomical changes in song control nuclei of the canary brain. *Science 214*, 1368–1370.

Oller, D. K. (2000). *The emergence of the speech capacity.* Mahwah, NJ: Lawrence Erlbaum Associates.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences 110*, 6318–632.

Oller, D. K. & Griebel, U. (2008a). Complexity and flexibility in infant vocal development and the earliest steps in the evolution of language. In D. K. Oller & U. Griebel (Eds.), *Evolution of communicative flexibility: Complexity, creativity and adaptability in human and animal communication.* (pp. 141–168). Cambridge, MA: MIT Press.

Oller, D. K., & Griebel, U. (2008b). The origins of syllabification in human infancy and in human evolution. In B. Davis & K. Zajdo (Eds.), E *Syllable development: The frame/content theory and beyond.* (pp. 29–62). Mahwah, NJ: Lawrence Erlbaum Associates.

Oller, D. K., & Griebel, U. (2014). On quantitative comparative research in communication and language evolution. *Biological Theory 9*(2). doi: 10.1007/s13752-014-0186-7

Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Topics in Cognitive Science* (Special issue, *New Frontiers in Language Evolution and Development*) *8*, 382–392.

Oller, D. K., Vohr, B. R., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Warlaumont, A. S., Griebel, U., & Buder, E. H. (2014). Infant vocalization and birdsong: An infrastructural view [Paper presented at the Society for Neuroscience satellite meeting]. *Birdsong 4: Rhythms and Clues from Neurons to Behavior.* Washington, DC.

Owings, D. H., & Morton, E. S. (1998). *Animal vocal communication.* Cambridge: Cambridge University Press.

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology 10*, 58–71.

Owren, M. J., Rendall, D., & Ryan, M. J. (2010). Redefining animal signaling: Influence versus information in communication. *Biological Philosophy 25*, 755–780.

Payne, R. B., & Mcvay, S. (1971). Songs of the humpback whale. *Science 173*, 583–597.

Payne, R. B., & Payne, L. L. (1997). Field observations, experimental design, and the time and place of learning bird songs. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development.* (pp. 57–84). Cambridge: Cambridge University Press.

Premack, D. (1971). Language in Chimpanzee? *Science 172*, 808–822.

Price, T. (2013). *Vocal communication within the genus chlorocebus: Insights into mechanisms of call production and call perception* (Ph.D thesis), Georg-August-Universität.

Rendall, D., Owren, M. J., & Ryan, M. J. (2009. What do animal signals mean? *Animal Behaviour 78*, 233–240.

Savage-Rumbaugh, E. S., & Macdonald, K. (1988). Deception and social manipulation in symbol-using apes. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence.* (pp. 224–237). Oxford: Clarendon Press.

Sayigh, L. S., Esch, H. C., Wells, R.S., & Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins. *Tursiops truncatus*. *Animal Behaviour 74*, 1631–1642.

Schagatay, E. (2011). Human breath-hold diving ability suggests a selective pressure for diving during human evolution. In M. Vaneechoutte, M. (Ed.), *Was man more aquatic in the past? fifty years after Alister Hardy—Waterside hypotheses of human evolution.* (pp. 120–147). Sharjah, UAE: Bentham Science.

Schusterman, R. J., Hanggi, E. B., & Gisiner, R. (1992). Acoustic signalling in mother-pup reunions, interspecies bonding, and affiliation by kinship in California sea lions (*Zalophus califonianus*). In J. Thomas, R. A. Kastelein, & A. Y. Supin (Eds.), *Marine mammal sensory systems.* (pp. 533–551). New York: Plenum 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.

Sinha, C. (2004). The evolution of language: From signals to symbols to system. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach.* (pp. 217–235). Cambridge, MA: MIT Press.

Snowdon, C. T. (2004). Social processes in the evolution of complex cognition and communication. In D. K. Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach.* (pp. 131–248). Cambridge, MA: MIT Press.

Stegmann, U. E. (2013). *Animal communication theory: Information and influence.* Cambridge: Cambridge University Press.

Stoeger, A. S., & Manger, P. (2014). Vocal learning in elephants: Neural bases and adaptive context. *Current Opinion in Neurobiology 28*, 101–107.

Stokoe, W. C. (1960). Sign language structure: An outline of the visual communication systems of the American Deaf [Occasional papers, Dept. of Anthropology, University of Buffalo], *Studies in Linguistics 8*, 1–78.

Sutton, D. (1979). Mechanisms underlying learned vocal control in primates. In H. D. Steklis & M. J. Raleigh (Eds.), *Neurobiology of social communication in primates: An evolutionary perspective.* (pp. 45–67). New York: Academic Press.

Suzuki, T. N. (2012). Referential mobbing calls elicit different predator-searching behaviours in Japanese great tits. *Animal Behaviour 84*, 53–57.

Tinbergen, N. (1951). *The study of instinct.* Oxford, Oxford University Press.

Tinbergen, N. (1952). Derived activities: Their causation, biological significance, origin and emancipation during evolution. *Quarterly Review of Biology 27*, 1–32.

Tobias, P. V. (2011). Revisiting water and hominin evolution. In M. Vaneechoutte (Ed.), *Was man more aquatic in the past? Fifty years after Alister Hardy—Waterside hypotheses of human evolution.* (pp. 3–15). Sharjah, UAE: Bentham Science.

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

Tyack, P. L. (2000). Dolphins whistle a signature tune. *Science 289*(5483), 1310–1311. **doi.org/10.1126/science.289.5483.1310**

Tyack, P. L. & Clark, C. W. 2000. Communication and acoustic behavior of dolphins and whales. In W. W. L. Au, A. N. Popper, & R. R. Fay (Eds.), *Hearing by Whales and Dolphins* (pp. 156–224). New York: Springer-Verlag.

Tyack, P. L. & Sayigh, L. (1997). Vocal learning in cetaceans. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development.* (pp. 208–233). New York: Cambridge University Press.

Verhaegen, M., Munro, S., Puech, P.-F., & Vaneechoutte, M. (2011). Early hominoids: Orthograde aquarboreals in flooded forests? In M. Vaneechoutte (Ed.), *Was man more aquatic in the past? Fifty years after Alister Hardy—Waterside hypotheses of human evolution.* (pp. 67–81). Sharjah, UAE: Bentham Science.

Verhaegen, M., Munro, S., Vaneechoutte, M., Bender-Oser, N., & Bender, R. (2007). The original econiche of the genus *Homo*: Open plain or waterside? In S. I. Munoz (Ed.), *Ecology Research Progress* (pp. 1–34). Hauppauge, NY: Nova Science Publishers.

Weiβ, B. M., Ladich, F., Spong, P., & Symonds, H. (2006). Vocal behavior of resident killer whale matrilines with newborn calves: The role of family signatures. *Journal of the Acoustical Society of America 119*, 627–635.

West, M. J., King, A. P., & Freeberg, T. M. (1997). Building a social agenda for the study of bird song. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development.* (pp. 41–56). Cambridge: Cambridge University Press.

Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology 21*, 195–205.

Winter, P., Ploog, D. W., & Latta, J. (1966). Vocal repertoire of the squirrel monkey (Saimiri sciureus), its analysis and significance. *Experimental Brain Research 1*, 359–384.

Wood, B. (2019). Beachcombing and coastal settlement: The long migration from South Africa to Patagonia—the greatest journey ever made. *Journal of Big History III*(4), 19–46.

Wrangham, R., Cheney, D., Seyfarth, R., & Sarmiento, E. (2009). Shallow-water habitats as sources of fallback foods for hominins. *American Journal of Physical Anthropology 140*, 630–642.

Ulrike Griebel

School of Communication Sciences and Disorders & Institute for Intelligent Systems, The University of Memphis; Konrad Lorenz Institute for Evolution and Cognition Research, Austria

D. Kimbrough Oller

School of Communication Sciences and Disorders & Institute for Intelligent Systems, The University of Memphis; Konrad Lorenz Institute for Evolution and Cognition Research, Austria