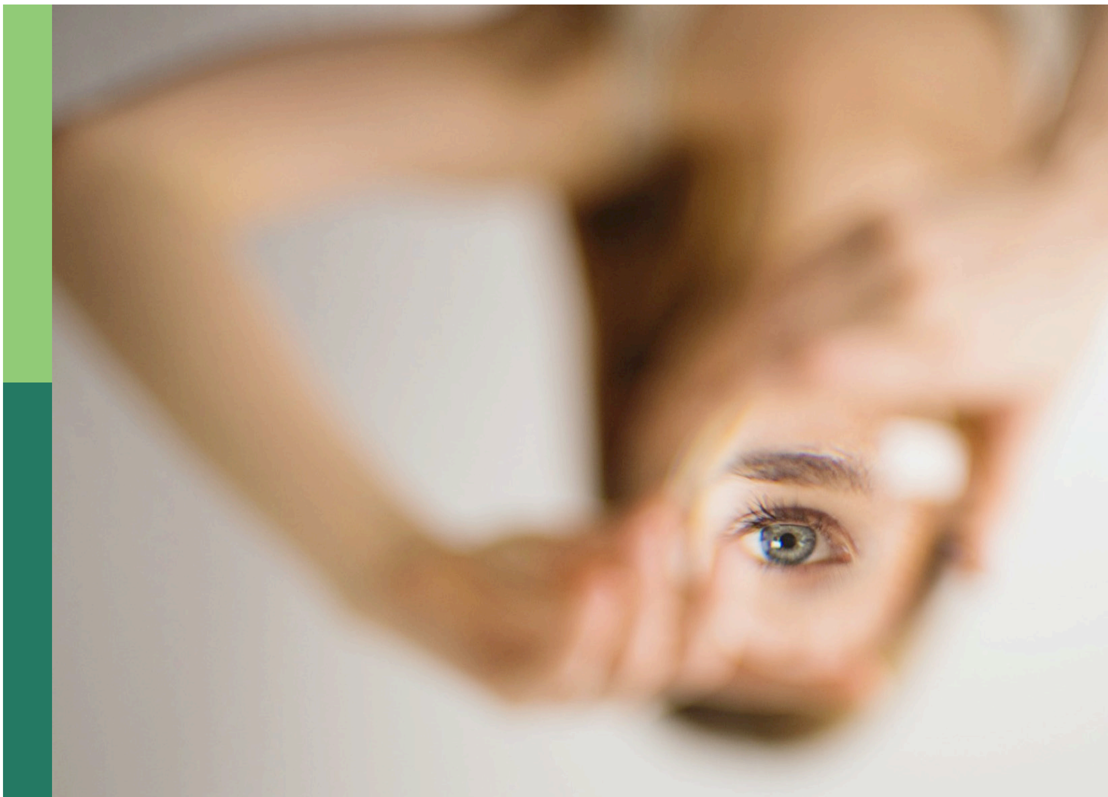
Research Topics

The quest for symbolic communication in non-human animals

Edited by
Ulrike Griebel, D. Kimbrough Oller
and Irene M. Pepperberg

Published in
Frontiers in Psychology



ISSN 1664-8714
ISBN 978-2-8325-5033-5
DOI 10.3389/978-2-8325-5033-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The quest for symbolic communication in non-human animals

Topic editors
Ulrike Griebel — University of Memphis, United States
D. Kimbrough Oller — University of Memphis, United States
Irene M. Pepperberg — Hunter College (CUNY), United States

Citation
Griebel, U., Oller, D. K., Pepperberg, I. M., eds. (2024). *The quest for symbolic communication in non-human animals*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-5033-5

Table of contents

MINI REVIEW
published: 12 November 2020
doi: 10.3389/fpsyg.2020.571678

Do Bats Have the Necessary Prerequisites for Symbolic Communication?

Mirjam Knörnschild[1,2,3] and Ahana A. Fernandez[1]*

[1] *Museum für Naturkunde, Leibniz-Institute for Evolution and Biodiversity Science, Berlin, Germany,* [2] *Animal Behavior Lab, Freie Universität, Berlin, Germany,* [3] *Smithsonian Tropical Research Institute, Ancón, Panama*

Training animals such as apes, gray parrots, or dolphins that communicate *via* arbitrary symbols with humans has revealed astonishing mental capacities that may have otherwise gone unnoticed. Albeit bats have not yet been trained to communicate *via* symbols with humans, we are convinced that some species, especially captive Pteropodid bats ("flying foxes"), show the potential to master this cognitive task. Here, we briefly review what is known about bats' cognitive skills that constitute relevant prerequisites for symbolic communication with humans. We focus on social learning in general, trainability by humans, associative learning from humans, imitation, vocal production learning and usage learning, and social knowledge. Moreover, we highlight potential training paradigms that could be used to elicit simple "symbolic" bat-human communication, i.e., training bats to select arbitrary symbols on a touchscreen to elicit a desired behavior of the human caregiver. Touchscreen-proficient bats could participate in cognition research, e.g., to study their numerical competence or categorical perception, to further elucidate how nonhuman animals learn and perceive the world.

Keywords: symbols, indexical communication, social learning, cognitive skills, touchscreen, training paradigm, bats, associative learning

OPEN ACCESS

Edited by:
Irene M. Pepperberg,
Harvard University, United States

Reviewed by:
Kirsten Bohn,
Johns Hopkins University,
United States
Darren Burke,
The University of Newcastle, Australia
Diana Reiss,
Hunter College (CUNY), United States

Correspondence:
Mirjam Knörnschild
mirjam.knoernschild@mfn.berlin;
mirjam.knoernschild@gmail.com

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *11 June 2020*
Accepted: *12 October 2020*
Published: *12 November 2020*

Citation:
Knörnschild M and
Fernandez AA (2020) Do Bats Have
the Necessary Prerequisites for
Symbolic Communication?
Front. Psychol. 11:571678.
doi: 10.3389/fpsyg.2020.571678

INTRODUCTION

Language is crucial to transmit information, share and accumulate knowledge across generations, and promote humans' cumulative culture (Tomasello, 2000; Herrmann et al., 2007; Fitch et al., 2010). Therefore, language drives and is driven by social cognition (Tomasello, 1992; Fitch et al., 2010). Besides a large set of physical cognitive skills, language particularly requires sociocognitive skills. Physical cognitive skills include memory, categorical perception and discrimination, perceptual processing, and recognition; and some researchers would also include general learning abilities such as fast mapping or associative learning as additional prerequisites (Gopnik et al., 1999; Vihman, 2014). Sociocognitive skills include, for example, social learning and theory of mind (Tomasello, 2003; Cheney and Seyfarth, 2007; Herrmann et al., 2007; Fitch et al., 2010). A remarkable form of social learning is our ability for imitation which plays a fundamental role in speech (or sign) acquisition (Oller, 1980; Petitto and Marentette, 1991; Vihman, 2014; Fitch, 2018). Infants acquire speech through imitation of the fundamental speech subunits, i.e., syllables, based on auditory input (Oller, 1980; Vihman et al., 1986). Whereas the ability of vocal production learning, i.e., the modification of one's own oral output based on social input, represents the mechanistic part of speech production, social knowledge is required to develop the semantic capacities of language (Tomasello, 1992, 2000; Fitch et al., 2010). The cognitive skills of joint attention, gaze responsiveness, and pointing pave the way for the developing the theory of mind in young infants (Carpenter and Tomasello, 1995; Gopnik et al., 1999; Tomasello, 2003). Joint attention, for example, is important for understanding others and enhances word learning (MacNamara, 1972; Gopnik et al., 1999; Tomasello, 2003). The development of these sociocognitive skills and, ultimately, language acquisition are shaped and promoted through social interaction (Tomasello, 1992; Kuhl, 2007; Goldstein and Schwade, 2010). Social feedback is also important for non-human vocal production learners (Goldstein and Schwade, 2010; Beecher, 2017; García, 2019), in particular, when learning non-species-specific vocalizations as the interaction in itself is already a form of communication (Pepperberg, 1992, 1994, 2002), or when learning to communicate *via* arbitrary symbols (Reiss and McCowan, 1993).

Language can be understood as a system of symbols whose elements (for example, words) can be arranged according to rules (through grammar) to create new meaningful units (such as sentences). Thus, the power of human symbolic communication is based upon the fact that the meaning of words can gain additional meaning through their relationship to other words, i.e., a sign-sign relationship (Sinha, 2004; Nieder, 2009). In contrast, non-human animal communication systems have indexical referential associations, i.e., they are based on a direct physical or temporal relation between sign-object or sign-event (Sinha, 2004; Nieder, 2009). The evolutionary transition from indexical communication in animals to symbolic communication in humans is considered to be associated with the emergence of language and symbolic thought (Deacon, 1998; Sinha, 2004; Nieder, 2009; Grouchy et al., 2016).

Even though only humans are thought to possess naturally occurring symbolic communication systems (i.e., natural languages, numerical systems), several other species such as apes, gray parrots, and dolphins can be trained to use symbols to express their needs/preferences when communicating with conspecifics (Fouts et al., 1984; Cianelli and Fouts, 1998; Pepperberg, 2009) or with humans (Gardner and Gardner, 1969; Herman et al., 1984; Schusterman and Krieger, 1984; Gisiner and Schusterman, 1992; Reiss and McCowan, 1993; Sevcik and Savage-Rumbaugh, 1994; Pepperberg, 2009). Symbolic communication between humans and animals can involve acoustic signals and speech (Herman et al., 1984; Pepperberg, 2009), gestures (Herman et al., 1984; Schusterman and Krieger, 1984, 1986), and technical interfaces such as TV monitors (Herman et al., 1990), interactive keyboards (Savage-Rumbaugh and Rumbaugh, 1978; Savage-Rumbaugh et al., 1980; Reiss and McCowan, 1993), or touchscreens (Nilsson et al., 2004; Amundin et al., 2008).

Training animals to communicate *via* arbitrary symbols has revealed astonishing mental capacities (Pepperberg, 1987, 2006; Boysen and Berntson, 1989; Reiss and McCowan, 1993; Savage-Rumbaugh and Fields, 2000; Kilian et al., 2003) which could have been overlooked if only the animals' naturally occurring communication signals had been decoded. When animals communicate with humans *via* learned arbitrary symbols, sign-object and sign-event relations are much more common than sign-sign relations (Sevcik and Savage-Rumbaugh, 1994; Pepperberg, 2009). Nevertheless, this simple "symbolic" communication is highly useful for understanding which cognitive prerequisites were necessary for the evolution of true symbolic communication, i.e., language in humans. Moreover, it allows for an in-depth investigation of species-specific mental capacities. Researchers documented, for example, cognitive skills such as numerical competence (Boysen and Berntson, 1989; Pepperberg, 2006), concept formation (Pepperberg, 1987), associative learning capabilities, and self-organized learning events (Reiss and McCowan, 1993).

Here, we want to give our perspective on the potential capability of bats to communicate with humans by using arbitrary symbols. Albeit bats have not yet been trained to communicate *via* symbols with humans, we are convinced for reasons that we outline below, that they show the potential to master this cognitive task. Bats are a very gregarious taxon comprising >1,400 extant species and exhibit a large spectrum of social systems with differing degrees of complexity (Wilkinson et al., 2019). Because taxonomic breadth is crucial for studying cognitive adaptations and achievements (Dukas, 2004), bats are an important taxon for comparative cognition research. Many bat species are long-lived (up to 30 years in the wild; Barclay and Harder, 2003) and most species either live in perennial stable groups (Wilkinson and Boughman, 1998) or have a social organization characterized by fission-fusion dynamics (Kerth, 2008). Both forms of temporal consistency in social interactions between group members pose different requirements on the cognitive abilities of the animals because they differ considerably in terms of relevant group size, frequency of repeated encounters, and consistency of social relationships.

Acoustic communication is one of the main channels for information transfer used by bats (Chaverri et al., 2018). In addition to echolocation (i.e., for navigation and foraging), different bat species possess diverse vocal repertoires and specific vocalization types which encode various information types such as emotional state (Bastian and Schmidt, 2008; Walter and Schnitzler, 2019) and identity information such as social group affiliation (Wilkinson and Boughman, 1998; Knörnschild et al., 2012), age (Jones et al., 1991; Fernandez and Knörnschild, 2017), and individual signatures (Carter et al., 2008; Chaverri et al., 2010). Vision and olfaction, the other two main sensory modalities in bats, are less well understood. Both phylogeny and species-specific dietary preferences influence bats' visual capabilities (**Figure 1**): whereas most Old Word fruit bats (*Pteropodidae*) rely almost exclusively on vision for orientation (Möhres and Kulzer, 1956), only some members of the genus *Rousettus* can use rudimentary echolocation based on tongue clicks (Grinnell and Hagiwara, 1972). Acoustics are of crucial importance to insectivorous bats which capture their prey *via* echolocation (Neuweiler, 1989). In contrast to insectivorous bats, nectarivorous and frugivorous bats have comparably larger eyes and a better

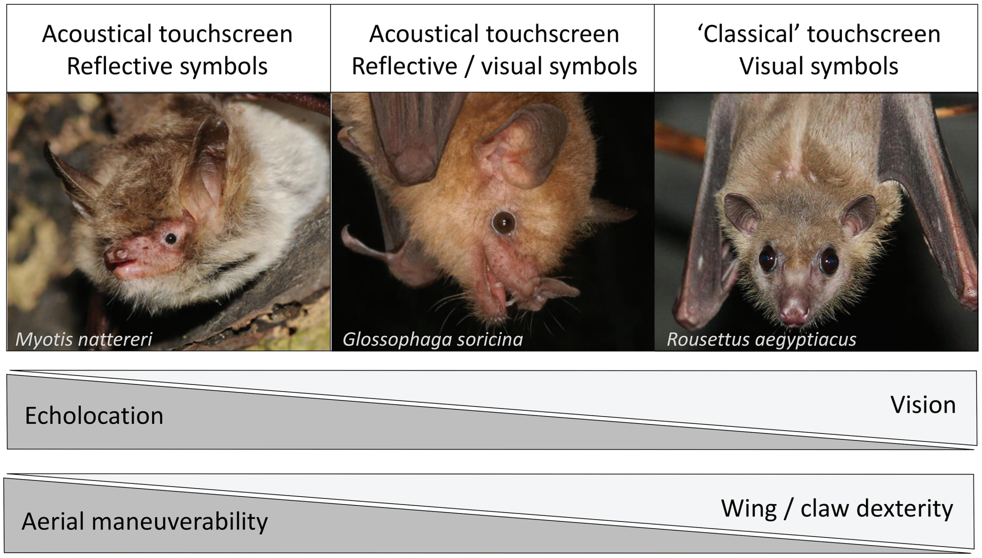


FIGURE 1 | Knowledge about species-specific strength and weaknesses in perception, maneuverability, and dexterity must inform the training paradigms for bat-human communication, e.g., with a touchscreen. Whereas most bats rely on echolocation to perceive the world, many species also use vision to a certain degree. For the Pteropodid bats ("flying foxes"), vision is the most important sense and only some members of the genus *Rousettus* can use rudimentary echolocation based on tongue clicks. Whereas Pteropodid bats reach comparatively high levels of dexterity with their wings and claws and often use them to manipulate objects, many non-Pteropodid bats do not. In turn, non-Pteropodid bats generally show greater aerial maneuverability than Pteropodid bats. Thus, visually oriented bats with high dexterity should be trained to use a "classical" touchscreen with visual symbols which they can approach by crawling/climbing whereas echoacoustically oriented bats with high aerial maneuverability should be trained to use a touchscreen with reflective symbols which they can activate with their sonar beam while hovering in front of it. If necessary, intermediate forms of these two extremes should be used to best accommodate a species' capabilities. The three depicted bat species represent the range of diverse species covered in the text: *Myotis nattereri*, an insectivorous gleaner (photo credit: Ján Svetlík), *Glossophaga soricina*, a nectarivorous flower-visiting bat (photo credit: Marco Tschapka), and *Rousettus aegyptiacus*, a frugivorous pteropodid (photo credit: Lithuanian Zoological Gardens).

vision (Zhao et al., 2009), even though they predominantly rely on echolocation as well, especially at short range-distances (Winter et al., 2005; Holland, 2007). Olfaction plays an important additional role for foraging Pteropodids and frugivorous or nectarivorous Neotropical bats (Korine and Kalko, 2005; Raghuram et al., 2009; Gonzalez-Terrazas et al., 2016). Olfactory signals are also important mediators for social communication (Safi and Kerth, 2003; Voigt et al., 2008). However, bat olfaction will not be discussed further as this sensory modality is not well suited for training paradigms discussed later.

In the following, we briefly review what cognitive skills that constitute relevant prerequisites for symbolic communication are already known to be present in bats. Furthermore, we highlight potential training paradigms which could be used to elicit simple "symbolic" bat-human communication, i.e., bats using learned arbitrary symbols to elicit a desired behavior of the human caregiver. We hope to highlight practical approaches for future studies on symbolic communication in bats.

SOCIAL LEARNING

Social learning occurs when animals learn from others that they observe or with whom they interact, for example, about foraging strategies or predator avoidance (Hoppitt and Laland, 2013). In bats, social learning is widespread and includes learning about roost- or food-related information as well as vocal production learning (reviewed in Wilkinson and Boughman, 1999; Wright, 2016). Learning from conspecifics has received much more attention than learning from heterospecific bats (Page and Bernal, 2020); the latter has been investigated in only a few species so far (Clarin et al., 2014; Patriquin et al., 2018). Moreover, the majority of studies demonstrated horizontal social learning, i.e., adults learning from adults, whereas vertical social learning, i.e., pups learning from adults, is currently understudied and yields both positive (Ripperger et al., 2019) and negative results (Rose et al., 2019). Although bats learn faster from other bats than from humans (Gaudet and Fenton, 1984; Clarin et al., 2014), humans can nevertheless elicit associative learning in bats and train them to perform specific actions (reviewed in Siemers and Page, 2009).

ASSOCIATIVE LEARNING

Bats readily learn to associate a particular cue with a specific outcome, either by themselves *via* trial-and-error learning or from others *via* social learning. Associative learning has been mainly demonstrated in a foraging context (reviewed in Wilkinson and Boughman, 1999; Wright, 2016). Bats can be trained to associate various novel cues with a food reward, e.g., light cues (Clarin et al., 2014), acoustic cues (Jones et al., 2013), echoacoustic, i.e., reflective cues (Simon et al., 2014), olfactory cues (Page et al., 2012), and visual cues (Manske and Schmidt, 1979). Gleaning bats, i.e., species that capture prey from substrates, seem to be especially well suited for food-related associative learning tasks (Siemers, 2001; Page and Ryan, 2006; Hulgard and Ratcliffe, 2014; Patriquin et al., 2018). Nectarivorous bats also exhibit strong associative learning in a foraging context and can be trained to discriminate fine-scale differences between sensory cues (von Helversen, 2004; Simon et al., 2006; Ross and Holderied, 2013) but they generally rely more on spatial cues than sensory cues (Thiele and Winter, 2005; Stich and Winter, 2006; Carter et al., 2010). Insectivorous bats can be trained to recognize 3-D objects as acoustic landmarks and associate them with safe passage through a net opening (Yu et al., 2019). In many species, learned associations are flexible and bats can be trained to reverse their initial associations (Page and Ryan, 2005; Clarin et al., 2013; Ross and Holderied, 2013). There is very little data on how long learned associations are remembered but current evidence suggests that bats have good short- and long-term memory (Ruczyński and Siemers, 2011; Page et al., 2012; Clarin et al., 2014; but see: Hernández-Montero et al., 2020). The above-mentioned examples used positive reinforcement but associative learning can also be negatively reinforced. Bats readily acquire taste aversions, e.g., by associating a novel acoustic cue with a noxious food reward (Bates and Fenton, 1990) or a novel flavor cue with an episode of toxicosis (Ratcliffe et al., 2003).

TRAINABILITY BY HUMANS

Various techniques can be applied to coax bats to participate in associative learning tasks (reviewed in Siemers and Page, 2009). Two important techniques for training bats are fading and shaping (Terrace, 1963; Shettleworth, 1998; Domjan, 2003). When fading, bats are gradually introduced to a new stimulus by altering the stimulus in small steps (Jones et al., 2013; Hemingway et al., 2020). Fading is especially important when studying reversal learning as it also allows the removal of a bat's response to a known stimulus (Page and Ryan, 2005, 2006). When shaping, the desired response of a bat is increasingly reinforced while non-desired responses are not reinforced (Barber et al., 2003). Shaping is also the technique of choice when training bats to perform certain behaviors on command. Captive Pteropodid bats ("flying foxes") can be readily trained for husbandry and vet checks; for instance, they can learn to follow a target, to unfold their wings in response to a hand signal, and to touch an item on demand (pers. communication Brian Pope, Lubee Bat Conservancy, USA). We are not aware that non-Pteropodid bats are being trained for husbandry and vet checks. However, temporarily captive non-Pteropodid bats can be trained to approach humans to retrieve a food reward, to wait on a perch until the onset of a stimulus, and to fly to a specific position when perceiving a stimulus (Tuttle, 2019).

IMITATION

Several bat species are capable of imitating conspecifics' actions. Naïve individuals have been shown to learn about novel foraging situations by paying close attention to knowledgeable conspecifics (*Eptesicus fuscus*: Wright et al., 2011; *Antrozous pallidus*: Bunkley and Barber, 2014). Imitation has also been shown in a communicative context, namely, when pups learn to sing by imitating the song of adult tutors (*Saccopteryx bilineata*: Knörnschild et al., 2010).

VOCAL PRODUCTION LEARNING AND USAGE LEARNING

Imitating new signals is one form of vocal production learning (VPL), modifying existing signals based on social influences is another (Janik and Slater, 1997, 2000). VPL *via* social modification has been shown for social calls (*Rousettus aegyptiacus*: Prat et al., 2015, 2017; Genzel et al., 2019; *Saccopteryx bilineata*: Knörnschild et al., 2012; *Phyllostomus discolor*: Esser and Schmidt, 1989; Esser, 1998; Lattenkamp et al., 2020; *P. hastatus*: Boughman, 1998) and echolocation calls (*Rhinolophus ferrumequinum*: Jones and Ransome, 1993; *Hipposideros terasensis*: Hiryu et al., 2006). In addition to VPL, vocal usage learning has been demonstrated by training temporarily isolated bats to vocalize in order to trigger a food reward (*P. discolor*: Lattenkamp et al., 2018). It is plausible that more bat species may have some degree of volitional control over their vocalizations but data are currently lacking.

SOCIAL KNOWLEDGE

Social knowledge describes the cognitive assessment of cues that communicate socially relevant information (Cheney et al., 1986). Whereas social knowledge mainly constitutes learning about others, such as their status or intentions, sociocognitive skills also facilitate the interpretation of signals or cues from others outside a social context (e.g., using gaze following to identify the location of food that a conspecific has hidden; Tomasello et al., 1998). In bats, social knowledge is severely understudied and most circumstantial evidence concerns comparatively simple sociocognitive skills such as the maintenance of dominance hierarchies (Neuweiler, 1969) or territorial interactions (Voigt and Streich, 2003). Advanced sociocognitive skills such as gaze following, joint attention, point following, and theory of mind are found to varying degrees in highly intelligent social species, such as primates and corvids, and also in domesticated species such as dogs; they can include heterospecific interactions, for example with humans (reviewed in Fitch et al., 2010). Evidence for heterospecific social knowledge in bats is currently limited to one study which demonstrated that captive born individuals of different bat species (*Pteropus pumilus*, *P. rodricensis*, and *P. conspicillatus*) are responsive to human pointing gestures (Hall et al., 2011): experimentally naïve bats readily utilize human pointing to find the location of concealed food in an object-choice task. The observed spontaneous point-following behavior suggests advanced sociocognitive skills in these bats. Interestingly, only captive born individuals were sensitive to human gestures; captive individuals born in the wild (*P. pumilus* and *P. vampyrus*) were not (Hall et al., 2011). It is possible that direct contact with humans early in ontogeny is necessary for bats to exhibit heterospecific point-following behavior.

DISCUSSION

There is conclusive evidence, albeit sometimes anecdotal, that different bat species possess several key prerequisites necessary for symbolic communication, most importantly associative learning and a general readiness to interact with and learn from caregivers in captivity. However, it is important to note that the ability for associative learning alone is not a guarantee that bats can transfer simple associations to more complex symbolic representations. What is missing so far is an experimental approach that actively combines these abilities to test if rudimentary symbolic bat-human communication can be achieved.

If attempted, we suggest making the task as easy as possible in both implementation and perception to facilitate the initial communication process. Training bats to communicate their choice between different preferred food items *via* arbitrary symbols would be a promising starting point to implement bat-human communication. Touchscreens are very promising tools for animal-human communication because they can be activated *via* fingers, snouts, tongues, beaks, and sonar beams, thus, making them accessible to a wide range of taxa (reviewed in Egelkamp and Ross, 2019). Bats would need (1) to learn to operate a touchscreen, (2) learn the association of a certain symbol with a specific food item, and (3) to use the symbol when communicating with a human *via* a touchscreen (**Figure 2**).

Accommodating species-specific differences in perception is crucial for the success of this endeavor (**Figure 1**). Visually oriented bats such as Pteropodids could be trained to use a touchscreen with visual symbols representing different preferred food items, as has been successfully done with primates (Savage-Rumbaugh, 1993). Echoacoustically oriented bats could be trained to use an acoustically activated touchscreen instead. This method, termed Echo Location Visualization and Interface System (ELVIS), has been developed for dolphins (Nilsson et al., 2004; Amundin et al., 2008) and allows them to use their sonar beam to "touch" and, thus, choose items on a screen, e.g., to communicate food preferences (Starkhammar et al., 2007). For bats, an acoustically activated touchscreen would ideally not depict visual symbols but reflective symbols (e.g., reliefs) to facilitate perception.

Even though bats are capable of vocal production and usage learning, we would advise against the use of acoustic symbols to facilitate bat-human communication. In contrast to certain songbirds, parrots, and dolphins, the imitation of heterospecific

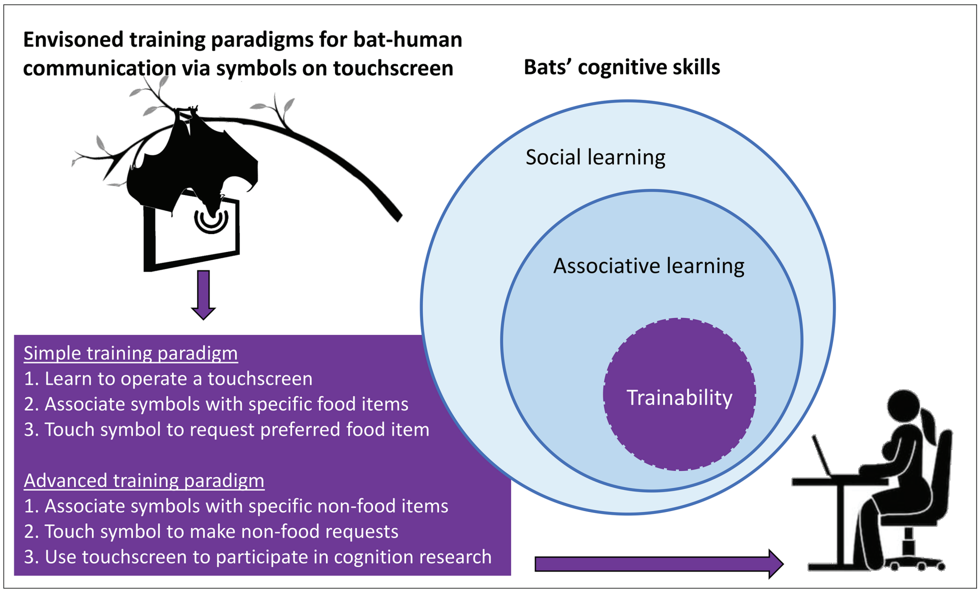


FIGURE 2 | Envisioned training paradigms for bat-human communication *via* symbols on a touchscreen. A set of different cognitive skills should enable bats to use a touchscreen, most importantly their capability of associative learning in a social context. We suggest focusing on visual or echoacoustic, i.e., reflective symbols on a touchscreen. A simple training paradigm requires bats to learn to operate a touchscreen by touching visual symbols with their snout (or reflective symbols with their sonar beam), to associate different symbols with specific food items, and to use these symbols to communicate which food item they prefer to receive (sign-object relation). An advanced training paradigm requires bats to associate different symbols with specific non-food items, e.g., caresses, access to toys, etc., and to use these symbols to communicate their preference (sign-event or sign-object relation). Touchscreen-proficient bats can participate in cognition research, e.g., to study their numerical competence or categorical perception.

sounds has never been demonstrated in bats. Because heterospecific vocal imitation is crucial for using novel sounds as symbols, we suggest focusing on visual or echoacoustic, i.e., reflective symbols for bat-human communication instead.

To conclude, bats are a promising taxon for future studies on symbolic communication with humans. Their willingness to interact with caregivers, associative learning abilities, and advanced (socio-)cognitive skills are important prerequisites to communicate successfully with humans. If bat-human communication about food requests could indeed be established, it would be an ideal stepping stone for a more advanced comparative cognition research, further elucidating how nonhuman animals think and learn.

AUTHOR CONTRIBUTIONS

MK and AF reviewed the literature and wrote the manuscript. Both the authors contributed to the article and approved the submitted version.

FUNDING

This study was financed by a Heisenberg Fellowship (DFG KN935 5-1) awarded to MK and an Elsa-Neuman Fellowship awarded to AF. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Programme (2014–2020)/ERC GA 804352.

ACKNOWLEDGMENTS

We are indebted to Irene Pepperberg, Diana Reiss, and two reviewers for their constructive feedback. We also thank Martina Nagy and Simon Ripperger for their comments on the manuscript and Merlin Tuttle and Brian Pope for sharing their knowledge on training paradigms. Moreover, we are grateful to Ján Svetlík, Marco Tschapka, and the Lithuanian Zoological Gardens for contributing bat photos. We acknowledge support by the Open Access Publication Initiative of Freie Universität Berlin.

REFERENCES

Amundin, M., Starkhammar, J., Evander, M., Almqvist, M., Lindstrom, K., and Persson, H. W. (2008). An echolocation visualization and interface system for dolphin research. *J. Acoust. Soc. Am.* 123, 1188–1194. doi: 10.1121/1.2828213

Barber, J. R., Razak, K. A., and Fuzessery, Z. M. (2003). Can two streams of auditory information be processed simultaneously? Evidence from the gleaning bat *Antrozous pallidus*. *J. Comp. Physiol. A* 189, 843–855. doi: 10.1007/s00359-003-0463-6

Barclay, R. M. R., and Harder, L. D. (2003). "Life histories of bats: life in the slow lane" in *Bat ecology*. eds. T. H. Kunz and B. Fenton (London: The University of Chicago Press, Ltd), 209–256.

Bastian, A., and Schmidt, S. (2008). Affect cues in vocalizations of the bat, *Megaderma lyra*, during agonistic interactions. *J. Acoust. Soc. Am.* 124, 598–608. doi: 10.1121/1.2924123

Bates, D. L., and Fenton, M. B. (1990). Aposematism or startle? Predators learn their responses to the defenses of prey. *Can. J. Zool.* 68, 49–52. doi: 10.1139/z90-009

Beecher, M. D. (2017). Birdsong learning as a social process. *Anim. Behav.* 124, 233–246. doi: 10.1016/j.anbehav.2016.09.001

Boughman, J. W. (1998). Vocal learning by greater spear-nosed bats. *Proc. R. Soc. Lond. B Biol. Sci.* 265, 227–233. doi: 10.1098/rspb.1998.0286

Boysen, S. T., and Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *J. Comp. Psychol.* 103, 23–31. doi: 10.1037/0735-7036.103.1.23

Bunkley, J. P., and Barber, J. R. (2014). An observation of apparent teaching behavior in the pallid bat, *Antrozous pallidus*. *West. N. Am. Nat.* 74, 249–252. doi: 10.3398/064.074.0213

Carpenter, M., and Tomasello, M. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Soc. Dev.* 4, 217–237. doi: 10.1111/j.1467-9507.1995.tb00063.x

Carter, G. G., Ratcliffe, J. M., and Galef, B. G. (2010). Flower bats (*Glossophaga soricina*) and fruit bats (*Carollia perspicillata*) rely on spatial cues over shapes and scents when relocating food. *PLoS One* 5:e10808. doi: 10.1371/journal.pone.0010808

Carter, G. G., Skowronski, M. D., Faure, P. A., and Fenton, B. (2008). Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim. Behav.* 76, 1343–1355. doi: 10.1016/j.anbehav.2008.04.023

Chaverri, G., Ancillotto, L., and Russo, D. (2018). Social communication in bats. *Biol. Rev.* 93, 1938–1954. doi: 10.1111/brv.12427

Chaverri, G., Gillam, E. H., and Vonhof, M. J. (2010). Social calls used by a leaf-roosting bat to signal location. *Biol. Lett.* 6, 441–444. doi: 10.1098/rsbl.2009.0964

Cheney, D. L., and Seyfarth, R. M. (2007). *Baboon metaphysics: The evolution of a social mind Chicago*. IL: University of Chicago Press.

Cheney, D., Seyfarth, R., and Smuts, B. (1986). Social relationships and social cognition in nonhuman primates. *Science* 234, 1361–1366. doi: 10.1126/science.3538419

Cianelli, S. N., and Fouts, R. S. (1998). Chimpanzee to chimpanzee American sign language. *J. Hum. Evol.* 13, 147–159. doi: 10.1007/BF02436502

Clarin, T. M. A., Borissov, I., Page, R. A., Ratcliffe, J. M., and Siemers, B. M. (2014). Social learning within and across species: information transfer in mouse-eared bats. *Can. J. Zool.* 92, 129–139. doi: 10.1139/cjz-2013-0211

Clarin, T. M. A., Ruczyński, I., Page, R. A., and Siemers, B. M. (2013). Foraging ecology predicts learning performance in insectivorous bats. *PLoS One* 8:e64823. doi: 10.1371/journal.pone.0064823

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. New York: W.W. Norton and Company Inc.

Domjan, M. (2003). *The principles of learning and memory*. *5th Edn*. Belmont, California: Wadsworth.

Dukas, R. (2004). Evolutionary biology of animal cognition. *Annu. Rev. Ecol. Evol. Syst.* 35, 347–374. doi: 10.1146/annurev.ecolsys.35.112202.130152

Egelkamp, C. L., and Ross, S. R. (2019). A review of zoo-based cognitive research using touchscreen interfaces. *Zoo Biol.* 38, 220–235. doi: 10.1002/zoo.21458

Esser, K. H. (1998). "Psychoacoustic studies in neotropical bats" in *Clinical psychoacoustics*. eds. S. Nielzén and O. Olsson (Sweden: Lund University Press), 45–59.

Esser, K. H., and Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, Phyllostomidae)—evidence for acoustic learning. *Ethology* 82, 156–168. doi: 10.1111/j.1439-0310.1989.tb00496.x

Fernandez, A. A., and Knörnschild, M. (2017). Isolation calls of the bat *Saccopteryx bilineata* encode multiple messages. *Anim. Behav. Cogn.* 4, 169–186. doi: 10.12966/abc.04.05.2017

Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annu. Rev. Linguist.* 4, 255–279. doi: 10.1146/annurev-linguistics-011817-045748

Fitch, W. T., Huber, L., and Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* 65, 795–814. doi: 10.1016/j.neuron.2010.03.011

Fouts, R. S., Fours, D. H., and Schoenfeld, D. (1984). Sign language conversational interaction between chimpanzees. *Sign Lang. Stud.* 42, 1–12.

García, N. C. (2019). What have we recently learned about song learning and social interactions? *Behav. Ecol.* 30, 1193–1195. doi: 10.1093/beheco/arz098

Gardner, R. A., and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 165, 664–672.

Gaudet, C. L., and Fenton, B. M. (1984). Observational learning in three species of insectivorous bats (Chiroptera). *Anim. Behav.* 32, 385–388. doi: 10.1016/S0003-3472(84)80273-0

Genzel, D., Desai, J., Paras, E., and Yartsev, M. M. (2019). Long-term and persistent vocal plasticity in adult bats. *Nat. Commun.* 10, 1–12. doi: 10.1038/s41467-019-11350-2

Gisiner, R., and Schusterman, R. J. (1992). Sequence, syntax, and semantics: responses of a language-trained sea lion (*Zalophus californianus*) to novel sign combinations. *J. Comp. Psychol.* 106, 78–91. doi: 10.1037/0735-7036.106.1.78

Goldstein, M. H., and Schwade, J. A. (2010). "From birds to words: perception of structure in social interactions guides vocal development and language learning" in *The Oxford handbook of developmental and comparative neuroscience*. eds. M. S. Blumberg, J. H. Freeman and S. R. Robinson (Oxford: Oxford University Press), 708–729.

Gonzalez-Terrazas, T. P., Martel, C., Milet-Pinheiro, P., Ayasse, M., Kalko, E. K., and Tschapka, M. (2016). Finding flowers in the dark: nectar-feeding bats integrate olfaction and echolocation while foraging for nectar. *R. Soc. Open Sci.* 3:160199. doi: 10.1098/rsos.160199

Gopnik, A., Meltzoff, A. N., and Kuhl, P. (1999). *The scientist in the crib: What early learning tells us about the mind*. New York: Harper Collins Publishers Inc.

Grinnell, A. D., and Hagiwara, S. (1972). Studies of auditory neurophysiology in non-echolocating bats, and adaptations for echolocation in one genus, *Rousettus*. *Z. Vergl. Physiol.* 76, 82–96. doi: 10.1007/BF00395501

Grouchy, P., D'Eleuterio, G. M., Christiansen, M. H., and Lipson, H. (2016). On the evolutionary origin of symbolic communication. *Sci. Rep.* 6, 1–9. doi: 10.1038/srep34615

Hall, N. J., Udell, M. A., Dorey, N. R., Walsh, A. L., and Wynne, C. D. (2011). Megachiropteran bats (*Pteropus*) utilize human referential stimuli to locate hidden food. *J. Comp. Psychol.* 125, 341–346. doi: 10.1037/a0023680

Hemingway, C. T., Ryan, M. J., and Page, R. A. (2020). State-dependent learning influences foraging behaviour in an acoustic predator. *Anim. Behav.* 163, 33–38. doi: 10.1016/j.anbehav.2020.02.004

Herman, L. M., Morrel-Samuels, P., and Pack, A. A. (1990). Bottlenosed dolphin and human recognition of veridical and degraded video displays of an artificial gestural language. *J. Exp. Psychol. Gen.* 119, 215–230. doi: 10.1037/0096-3445.119.2.215

Herman, L. M., Richards, D. G., and Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition* 16, 129–219. doi: 10.1016/0010-0277(84)90003-9

Hernández-Montero, J. R., Schöner, C. R., and Kerth, G. (2020). No evidence for memory retention of a learned association between a cue and roost quality after hibernation in free-ranging bats. *Ethology* 126, 761–771. doi: 10.1111/eth.13029

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366. doi: 10.1126/science.1146282

Hiryu, S., Katsura, K., Nagato, T., Yamazaki, H., Lin, L. -K., Watanabe, Y., et al. (2006). Intra-individual variation in the vocalized frequency of the Taiwanese leaf-nosed bat, *Hipposideros terasensis*, influenced by conspecific colony members. *J. Comp. Physiol. A* 192, 807–815. doi: 10.1007/s00359-006-0118-5

Holland, R. A. (2007). Orientation and navigation in bats: known unknowns or unknown unknowns? *Behav. Ecol. Sociobiol.* 61, 653–660. doi: 10.1007/s00265-006-0297-7

Hoppitt, W., and Laland, K. N. (2013). *Social learning: An introduction to mechanisms, methods, and models*. Princeton: Princeton University Press.

Hulgard, K., and Ratcliffe, J. M. (2014). Niche-specific cognitive strategies: object memory interferes with spatial memory in the predatory bat *Myotis nattereri*. *J. Exp. Biol.* 217, 3293–3300. doi: 10.1242/jeb.103549

Janik, V. M., and Slater, P. J. B. (1997). "Vocal learning in mammals" in *Advances in the study of behavior*. eds. P. J. B. Slater, J. S. Rosenblatt, C. T. Snowdon and M. Milinski (USA: Academic Press), 59–100.

Janik, V. M., and Slater, P. J. (2000). The different roles of social learning in vocal communication. *Anim. Behav.* 60, 1–11. doi: 10.1006/anbe.2000.1410

Jones, G., Hughes, P. M., and Rayner, J. M. V. (1991). The development of vocalizations in *Pipistrellus pipistrellus* (Chiroptera: Vespertilionidae) during post-natal growth and the maintenance of individual vocal signatures. *J. Zool.* 225, 71–84. doi: 10.1111/j.1469-7998.1991.tb03802.x

Jones, G., and Ransome, R. D. (1993). Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proc. R. Soc. Lond. B Biol. Sci.* 252, 125–128. doi: 10.1098/rspb.1993.0055

Jones, P. L., Ryan, M. J., Flores, V., and Page, R. A. (2013). When to approach novel prey cues? Social learning strategies in frog-eating bats. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20132330. doi: 10.1098/rspb.2013.2330

Kerth, G. (2008). Causes and consequences of sociality in bats. *Bioscience* 58, 737–746. doi: 10.1641/B580810

Kilian, A., Yaman, S., von Fersen, L., and Güntürkün, O. (2003). A bottlenose dolphin discriminates visual stimuli differing in numerosity. *Anim. Learn. Behav.* 31, 133–142. doi: 10.3758/BF03195976

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., and von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biol. Lett.* 6, 156–159. doi: 10.1098/rsbl.2009.0685

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., and von Helversen, O. (2012). Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Anim. Behav.* 84, 761–769. doi: 10.1016/j.anbehav.2012.06.029

Korine, C., and Kalko, E. K. (2005). Fruit detection and discrimination by small fruit-eating bats (Phyllostomidae): echolocation call design and olfaction. *Behav. Ecol. Sociobiol.* 59, 12–23. doi: 10.1007/s00265-005-0003-1

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Dev. Sci.* 10, 110–120. doi: 10.1111/j.1467-7687.2007.00572.x

Lattenkamp, E. Z., Vernes, S. C., and Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *J. Exp. Biol.* 221:jeb180729. doi: 10.1242/jeb.180729

Lattenkamp, E. Z., Vernes, S. C., and Wiegrebe, L. (2020). Vocal production learning in the pale spear-nosed bat, *Phyllostomus discolor*. *Biol. Lett.* 16:20190928. doi: 10.1098/rsbl.2019.0928

MacNamara, J. (1972). Cognitive basis of language learning in infants. *Psychol. Rev.* 79, 1–13. doi: 10.1037/h0031901

Manske, U., and Schmidt, U. (1979). Untersuchungen zur optischen Musterunterscheidung bei der Vampirfledermaus, *Desmodus rotundus*. *Z. Tierpsychol.* 49, 120–131. doi: 10.1111/j.1439-0310.1979.tb00280.x

Möhres, F. P., and Kulzer, E. (1956). Über die Orientierung der Flughunde (Chiroptera-Pteropodidae). *Z. Vergl. Physiol.* 38, 1–29. doi: 10.1007/BF00338621

Neuweiler, G. (1969). Verhaltensbeobachtungen an einer indischen Flughundkolonie (*Pteropus g. giganteus* Brünn). *Z. Tierpsychol.* 26, 166–199. doi: 10.1111/j.1439-0310.1969.tb01944.x

Neuweiler, G. (1989). Foraging ecology and audition in echolocating bats. *Trends Ecol. Evol.* 4, 160–166. doi: 10.1016/0169-5347(89)90120-1

Nieder, A. (2009). Prefrontal cortex and the evolution of symbolic reference. *Curr. Opin. Neurobiol.* 19, 99–108. doi: 10.1016/j.conb.2009.04.008

Nilsson, M., Lindström, K., Amundin, M., and Persson, H. W. (2004). Echolocation and visualization interface system ELVIS. Abstracts from the Symposium on the Occasion of the 50th Anniversary of Echocardiography, Sweden, 24, 169–178.

Oller, D. K. (1980). "The emergence of the sounds of speech in infancy" in *Child phonology*. eds. G. H. Yeni-Komshian, J. F. Kavanagh and C. A. Ferguson (New York: Academic Press), 93–112.

Page, R. A., and Bernal, X. E. (2020). The challenge of detecting prey: private and social information use in predatory bats. *Funct. Ecol.* 34, 344–363. doi: 10.1111/1365-2435.13439

Page, R. A., and Ryan, M. J. (2005). Flexibility in assessment of prey cues: frog-eating bats and frog calls. *Proc. R. Soc. Lond. B Biol. Sci.* 272, 841–847. doi: 10.1098/rspb.2004.2998

Page, R. A., and Ryan, M. J. (2006). Social transmission of novel foraging behavior in bats: frog calls and their referents. *Curr. Biol.* 16, 1201–1205. doi: 10.1016/j.cub.2006.04.038

Page, R. A., von Merten, S., and Siemers, B. M. (2012). Associative memory or algorithmic search: a comparative study on learning strategies of bats and shrews. *Anim. Cogn.* 15, 495–504. doi: 10.1007/s10071-012-0474-1

Patriquin, K. J., Kohles, J. E., Page, R. A., and Ratcliffe, J. M. (2018). Bats without borders: predators learn novel prey cues from other predatory species. *Sci. Adv.* 4:eaaq0579. doi: 10.1126/sciadv.aaq0579

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. *Anim. Learn. Behav.* 15, 423–432. doi: 10.3758/BF03205051

Pepperberg, I. M. (1992). A review of the effects of social interaction on vocal learning in African grey parrots (*Psittacus erithacus*). *Netherlands J. Zool.* 43, 104–124. doi: 10.1163/156854293X00241

Pepperberg, I. M. (1994). Vocal learning in grey parrots (*Psittacus erithacus*): effects of social interaction, reference, and context. *Auk* 111, 300–313. doi: 10.2307/4088595

Pepperberg, I. M. (2002). Cognitive and communicative abilities of grey parrots. *Curr. Dir. Psychol. Sci.* 11, 83–87. doi: 10.1111/1467-8721.00174

Pepperberg, I. M. (2006). Grey parrot numerical competence: a review. *Anim. Cogn.* 9, 377–391. doi: 10.1007/s10071-006-0034-7

Pepperberg, I. M. (2009). *The Alex studies: Cognitive and communicative abilities of grey parrots*. Cambridge, Massachusetts: Harvard University Press.

Petitto, L. A., and Marentette, P. F. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251, 1493–1496. doi: 10.1126/science.2006424

Prat, Y., Azoulay, L., Dor, R., and Yovel, Y. (2017). Crowd vocal learning induces vocal dialects in bats: playback of conspecifics shapes fundamental frequency usage by pups. *PLoS Biol.* 15:e2002556. doi: 10.1371/journal.pbio.2002556

Prat, Y., Taub, M., and Yovel, Y. (2015). Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats. *Sci. Adv.* 1:e1500019. doi: 10.1126/sciadv.1500019

Raghuram, H., Thangadurai, C., Gopukumar, N., Nathar, K., and Sripathi, K. (2009). The role of olfaction and vision in the foraging behaviour of an echolocating megachiropteran fruit bat, *Rousettus leschenaulti* (Pteropodidae). *Mamm. Biol.* 74, 9–14. doi: 10.1016/j.mambio.2008.02.008

Ratcliffe, J. M., Fenton, M. B., and Galef, B. G. Jr. (2003). An exception to the rule: common vampire bats do not learn taste aversions. *Anim. Behav.* 65, 385–389. doi: 10.1006/anbe.2003.2059

Reiss, D., and McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *J. Comp. Psychol.* 107, 301–312. doi: 10.1037/0735-7036.107.3.301

Ripperger, S., Günther, L., Wieser, H., Duda, N., Hierold, M., Cassens, B., et al. (2019). Proximity sensors on common noctule bats reveal evidence that mothers guide juveniles to roosts but not food. *Biol. Lett.* 15:20180884. doi: 10.1098/rsbl.2018.0884

Rose, A., Wöhl, S., Bechler, J., Tschapka, M., and Knörnschild, M. (2019). Maternal mouth-to-mouth feeding behaviour in flower-visiting bats, but no experimental evidence for transmitted dietary preferences. *Behav. Process.* 165, 29–35. doi: 10.1016/j.beproc.2019.06.001

Ross, G., and Holderied, M. W. (2013). "Learning and memory in bats: a case study on object discrimination in flower-visiting bats" in *Bat evolution, ecology, and conservation*. eds. R. A. Adams and S. C. Pedersen (New York, NY: Springer), 207–224.

Ruczyński, I., and Siemers, B. M. (2011). Hibernation does not affect memory retention in bats. *Biol. Lett.* 7, 153–155. doi: 10.1098/rsbl.2010.0585

Safi, K., and Kerth, G. (2003). Secretions of the interaural gland contain information about individuality and colony membership in the Bechstein's bat. *Anim. Behav.* 65, 363–369. doi: 10.1006/anbe.2003.2067

Savage-Rumbaugh, E. S. (1993). "Language learnability in man, ape, and dolphin" in *Language and communication*. eds. H. L. Roitblat, L. M. Herman and P. E. Nachtigall (New York: Lawrence Erlbaum Associates, Inc), 456–469.

Savage-Rumbaugh, E. S., and Fields, W. M. (2000). Linguistic, cultural and cognitive capacities of bonobos (*Pan paniscus*). *Cult. Psychol.* 6, 131–153. doi: 10.1177/1354067X0062003

Savage-Rumbaugh, E. S., and Rumbaugh, D. M. (1978). Symbolization, language, and chimpanzees: a theoretical reevaluation based on initial language acquisition processes in four young *Pan troglodytes*. *Brain Lang.* 6, 265–300. doi: 10.1016/0093-934X(78)90063-9

Savage-Rumbaugh, E. S., Rumbaugh, D. M., and Boysen, S. (1980). "Linguistically mediated tool use and exchange by chimpanzees (*Pan troglodytes*)" in *Speaking of apes*. eds. T. A. Sebeok and J. Umiker-Sebeok (Boston, MA: Springer), 353–383.

Schusterman, R. J., and Krieger, K. (1984). California sea lions are capable of semantic comprehension. *Psychol. Rec.* 34, 3–23. doi: 10.1007/BF03394849

Schusterman, R. J., and Krieger, K. (1986). Artificial language comprehension and size transposition by a California sea lion (*Zalophus californianus*). *J. Comp. Psychol.* 100, 348–355. doi: 10.1037/0735-7036.100.4.348

Sevcik, R. A., and Savage-Rumbaugh, E. S. (1994). Language comprehension and use by great apes. *Lang. Commun.* 14, 37–58. doi: 10.1016/0271-5309(94)90019-1

Shettleworth, S. J. (1998). *Cognition, evolution, and behavior*. New York: Oxford University Press.

Siemers, B. M. (2001). Finding prey by associative learning in gleaning bats: experiments with a Natterer's bat *Myotis nattereri*. *Acta Chiropterol.* 3, 211–215.

Siemers, B. M., and Page, R. A. (2009). "Behavioral studies of bats in captivity: methodology, training and experimental design" in *Ecological and behavioral methods for the study of bats*. eds. T. H. Kunz and S. Parsons (Baltimore: Johns Hopkins University Press), 373–392.

Simon, R., Holderied, M. W., and Von Helversen, O. (2006). Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat. *J. Exp. Biol.* 209, 3599–3609. doi: 10.1242/jeb.02398

Simon, R., Knörnschild, M., Tschapka, M., Schneider, A., Passauer, N., and von Helversen, O. (2014). Biosonar resolving power: echo-acoustic perception of surface structures in the submillimeter range. *Front. Physiol.* 5:64. doi: 10.3389/fphys.2014.00064

Sinha, C. (2004). "The evolution of language: from signals to symbols to systems" in *Evolution of communication systems, a comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge: MIT Press), 217–235.

Starkhammar, J., Amundin, M., Olsén, H., Almqvist, M., Lindström, K., and Persson, H. W. (2007). "Acoustic touch screen for dolphins, first application of ELVIS – an echo-location visualization and interface system" in *Proceedings from the 4th International Conference on Bio-Acoustics*, 29; April 10-12, 2007; UK, 63–68.

Stich, K. P., and Winter, Y. (2006). Lack of generalization of object discrimination between spatial contexts by a bat. *J. Exp. Biol.* 209, 4802–4808. doi: 10.1242/jeb.02574

Terrace, H. S. (1963). Discrimination learning with and without "errors." *J. Exp. Anal. Behav.* 6, 1–27. doi: 10.1901/jeab.1963.6-1

Thiele, J., and Winter, Y. (2005). Hierarchical strategy for relocating food targets in flower bats: spatial memory versus cue-directed search. *Anim. Behav.* 69, 315–327. doi: 10.1016/j.anbehav.2004.05.012

Tomasello, M. (1992). The social bases of language acquisition. *Soc. Dev.* 1, 67–87. doi: 10.1111/j.1467-9507.1992.tb00135.x

Tomasello, M. (2000). Culture and cognitive development. *Curr. Dir. Psychol. Sci.* 9, 37–40. doi: 10.1111/1467-8721.00056

Tomasello, M. (2003). "The key is social cognition" in *Language in mind: Advances in the study of language and thought*. eds. D. Gentner and S. Goldin-Meadow (Cambridge: MIT Press), 47–57.

Tomasello, M., Call, J., and Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Anim. Behav.* 55, 1063–1069. doi: 10.1006/anbe.1997.0636

Tuttle, M. (2019). How I photograph bats. Merlin Tuttle's Bat Conservation, Resources. Available at: https://www.merlintuttle.org/resources/how-i-photograph-bats/ (Accessed October 31, 2020).

Vihman, M. M. (2014). *Phonological development: The first two years*. USA: John Wiley and Sons.

Vihman, M. M., Ferguson, C. A., and Elbert, M. (1986). Phonological development from babbling to speech: common tendencies and individual-differences. *Appl. Psycholinguist.* 7, 3–40. doi: 10.1017/S0142716400007165

Voigt, C. C., Behr, O., Caspers, B., von Helversen, O., Knörnschild, M., Mayer, F., et al. (2008). Songs, scents, and senses: sexual selection in the greater sac-winged bat, *Saccopteryx bilineata*. *J. Mammal.* 89, 1401–1410. doi: 10.1644/08-MAMM-S-060.1

Voigt, C. C., and Streich, W. J. (2003). Queuing for harem access in colonies of the greater sac-winged bat. *Anim. Behav.* 65, 149–156. doi: 10.1006/anbe.2002.2031

von Helversen, D. (2004). Object classification by echolocation in nectar feeding bats: size-independent generalization of shape. *J. Comp. Physiol. A* 190, 515–521. doi: 10.1007/s00359-004-0492-9

Walter, M. H., and Schnitzler, H. U. (2019). Spectral call features provide information about the aggression level of greater mouse-eared bats (*Myotis myotis*) during agonistic interactions. *Bioacoustics* 28, 1–25. doi: 10.1080/09524622.2017.1359798

Wilkinson, G. S., and Boughman, J. (1998). Social calls coordinate foraging in greater spear-nosed bats. *Anim. Behav.* 55, 337–350. doi: 10.1006/anbe.1997.0557

Wilkinson, G. S., and Boughman, J. W. (1999). "Social influences on foraging in bats" in *Mammalian social learning: comparative and ecological perspectives*. eds. H. O. Box and K. R. Gibson (Cambridge: Cambridge University Press), 189–204.

Wilkinson, G. S., Carter, G., Bohn, K. M., Caspers, B., Chaverri, G., Farine, D., et al. (2019). Kinship, association, and social complexity in bats. *Behav. Ecol. Sociobiol.* 73:7. doi: 10.1007/s00265-018-2608-1

Winter, Y., von Merten, S., and Kleindienst, H. U. (2005). Visual landmark orientation by flying bats at a large-scale touch and walk screen for bats, birds and rodents. *J. Neurosci. Methods* 141, 283–290. doi: 10.1016/j.jneumeth.2004.07.002

Wright, G. S. (2016). "Social learning and information transfer in bats: conspecific influence regarding roosts, calls, and food" in *Sociality in bats*. ed. J. Ortega (Switzerland: Springer Press), 211–230.

Wright, G. S., Wilkinson, G. S., and Moss, C. F. (2011). Social learning of a novel foraging task by big brown bats, *Eptesicus fuscus*. *Anim. Behav.* 82, 1075–1083. doi: 10.1016/j.anbehav.2011.07.044

Yu, C., Luo, J., Wohlgemuth, M., and Moss, C. F. (2019). Echolocating bats inspect and discriminate landmark features to guide navigation. *J. Exp. Biol.* 222:jeb191965. doi: 10.1242/jeb.191965

Zhao, H., Rossiter, S. J., Teeling, E. C., Li, C., Cotton, J. A., and Zhang, S. (2009). The evolution of color vision in nocturnal mammals. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8980–8985. doi: 10.1073/pnas.0813201106

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

HYPOTHESIS AND THEORY
published: 26 January 2021
doi: 10.3389/fpsyg.2020.626138

Functionally Flexible Signaling and the Origin of Language

D. Kimbrough Oller[1,2,3]* *and Ulrike Griebel*[1,2,3]

[1]School of Communication Sciences and Disorders, The University of Memphis, Memphis, TN, United States, [2]Institute for Intelligent Systems, The University of Memphis, Memphis, TN, United States, [3]Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria

At the earliest break of ancient hominins from their primate relatives in vocal communication, we propose a selection pressure on vocal fitness signaling by hominin infants. Exploratory vocalizations, not tied to expression of distress or immediate need, could have helped persuade parents of the wellness and viability of the infants who produced them. We hypothesize that hominin parents invested more in infants who produced such signals of fitness plentifully, neglecting or abandoning them less often than infants who produced the sounds less frequently. Selection for such exploratory vocalization provided a critically important inclination and capability relevant to language, we reason, because the system that encouraged spontaneous vocalization also made vocalization functionally flexible to an extent that has not been observed in any other animal. Although this vocal flexibility did not by itself create language, it provided an essential foundation upon which language would evolve through a variety of additional steps. In evaluating this speculation, we consider presumable barriers to evolving language that are thought to be implications of Darwinian Theory. It has been claimed that communication always involves sender self-interest and that self-interest leads to deceit, which is countered through clever detection by receivers. The constant battle of senders and receivers has been thought to pose an insuperable challenge to honest communication, which has been viewed as a requirement of language. To make communication honest, it has been proposed that stable signaling requires costly handicaps for the sender, and since language cannot entail high cost, the reasoning has suggested an insurmountable obstacle to the evolution of language. We think this presumed honesty barrier is an illusion that can be revealed by recognition of the fact that language is not inherently honest and in light of the distinction between illocutionary force and semantics. Our paper also considers barriers to the evolution of language (not having to do with honesty) that we think may have actually played important roles in preventing species other than humans from evolving language.

Keywords: vocal development, honest signaling, origin of language, evolution of language, babbling, vocal learning, comparative psychology

OPEN ACCESS

Edited by:
Jan Antfolk,
Åbo Akademi University, Finland

Reviewed by:
Julia Fischer,
Deutsches Primatenzentrum, Germany
Judith M. Burkart,
University of Zurich, Switzerland

****Correspondence:***
D. Kimbrough Oller
koller@memphis.edu

Specialty section:
This article was submitted to Evolutionary Psychology, a section of the journal Frontiers in Psychology

Received: *04 November 2020*
Accepted: *30 December 2020*
Published: *26 January 2021*

Citation:
Oller DK and Griebel U (2021) Functionally Flexible Signaling and the Origin of Language. Front. Psychol. 11:626138. doi: 10.3389/fpsyg.2020.626138

OVERVIEW

A key goal of our research is to discover the most fundamental vocal capabilities and inclinations upon which language was founded, long before the first word was spoken. Further, we seek to posit evolutionary pressures that may have selected for these capabilities, a task that requires positing advantages, which could not have involved the advantages of language, which did not exist at that point. An initial step that may have moved hominins beyond the primate communicative background is *vocal functional flexibility* (VFF). We have long argued that the natural laboratory of human vocal development provides key evidence relevant to the search for origins of language (Oller, 2000; Griebel and Oller, 2008). VFF is seen from the first month of life in human infant "protophones" (Papaeliou et al., 2002; Scheiner et al., 2002; Oller et al., 2013; Jhang and Oller, 2017), the precursors to speech, including categories termed squeals, vocants (vowel-like sounds), and growls. The protophones serve different functions on different occasions. All protophone types are usually produced playfully or exploratorily with no obvious social intention or social directivity and with neutral facial affect (Long et al., 2020). But the same sounds are also produced on different occasions with positive or negative facial affect, suggesting, for example, exultation or complaint. For example, a squeal sound can be used (1) on one occasion with a big smile portraying apparent exultation, (2) on another occasion with an intense grimace, making an apparent complaint, even suggesting the infant is about to start crying, and (3) on yet another occasion, when the same infant is alone and playing quietly, with a neutral facial expression and no apparent social intent, merely exploring the sound.

Vocal functional flexibility is present throughout human life, since every linguistic signal must be functionally flexible. Any word, for example, must be able to serve a wide variety of different functions ("illocutionary forces," see below) on different occasions of use. We must even be able to pronounce any word just for the interest of doing it. The two facts (1) that VFF is present from the first month of human life, and (2) that VFF is a foundational requirement of vocal language, suggest that one of the first evolutionary steps that differentiated ancient hominins from their primate relatives in communicative capabilities may have been VFF.

In other primates, vocal flexibility is far more limited because their vocal signals appear to be required to have *particular beneficial effects in the here and now* – later effects are of course possible, but not the focus of the pressures that selected the signals. In this paper, we propose an evolutionary scenario where hominin infant fitness signaling through vocalizations with VFF could have been naturally selected. Importantly these vocalizations would have often had no necessary *immediate* communicatively generated benefits to the infant, just as is the case with modern human protophones. The primary benefits could occur later, when caregivers could invest in infant welfare based on a cumulative conscious or unconscious recollection of the infant fitness signals. The key point is that vocal signals of infant hominins, in this scenario, were selected in a way that left them free of immediate socio-functional requirements. From this platform of infant vocalization and parental awareness of it, we propose that natural selection of infants who showed vocal fitness signaling could have instigated selection of steadily increasing VFF in hominins, thus forming a foundation for and moving them in the direction of language. Subsequent steps built upon the foundation of VFF were, in accord with our proposal, necessary to establish symbolic content in signaling.

Our paper will consider the barrier to the evolution of language that has most often been proposed. The contention is that human language constitutes "honest signaling" (Fitch, 2004) and that because communication is inherently selfish and therefore inclined to deception (Dawkins and Krebs, 1978), language evolution is problematical. In a rebuttal of this line of reasoning, we shall argue that language is not in fact inherently honest, and we shall elucidate this fact by unpacking the distinction between illocutionary force and semantic content in communication (Austin, 1962), a distinction that also helps to illustrate and clarify the nature of functional flexibility. In Appendix A of the **Supplementary Material**, we offer additional reasons to reject the honest signaling argument. In Appendix B, we supply additional thoughts about a strategy for research on the origin of language along with possible foundations of language that can be seen in evolved communication signals of other species.

THE CRITICAL NATURE OF VOCAL FUNCTIONAL FLEXIBILITY IN LANGUAGE

Language is a capability and an inclination that evolved in ancient humans but must be developed within each individual. The emphasis on *inclination* is important because humans use language copiously, imaginatively, and often frivolously, sometimes with no social purpose but just for the pleasure of toying with language itself. In addition, as indicated above, from the first month of life, human infants produce protophones, not bound by any particular emotional state. In fact protophones are produced most commonly in apparent comfort and lack of immediate social goals (Oller et al., 2013; Jhang and Oller, 2017). Even when infants are alone in a room and comfortable, all-day recordings show that protophone production is common, yielding 3–4 utterances per minute (Oller et al., 2019a), and similar rates are observed for infants in the presence of a mother who, for example, is reading silently (Iyer et al., 2016). It is important, however, to emphasize that all the types of protophones that have been recognized as pertaining to the common infant repertoire are also produced in varying states of positive or negative emotion on different occasions, suggesting the protophones can indeed be used to express states with immediate communicative import, e.g., intended to solicit immediate attention from the caregiver.

Counts of protophones based on all-day recordings of infants in their homes show a huge rate, ~5 per minute during wakefulness, ~3,500 per day (Oller et al., 2019a), which is 5–10 times higher than the rate of crying even in the first month. The research shows that protophone production at high rates occurs from as soon as human infants can breathe on their own, as illustrated through all-day recordings of prematurely-born infants still in neonatal intensive care. Evidence of the robustness of the tendency to produce protophones copiously has been observed in American and European infants that have been studied longitudinally for many years (Stark, 1981; Elbers, 1982; Koopmans-van Beinum and van der Stelt, 1986; Stoel-Gammon, 1992), across infants with very different levels of socio-economic status (Eilers et al., 1993, 1997; Oller et al., 1995), across infants with very different languages in the home (Oller and Eilers, 1982; Holmgren et al., 1986; Lee et al., 2017), and even across infants who are later diagnosed with a wide variety of communication disorders (Oller and Eilers, 1988; Vinter, 1994; Masataka, 2001; Patten et al., 2014; Nyman and Lohmander, 2018). This seemingly obsessive human vocal tendency does not subside later in life, with all-day recordings suggesting human adults speaking English produce on the order of 16,000 words per day (Mehl et al., 2007).

There is abundant evidence that this human inclination to speak is endogenous. Consider how often we adults talk to ourselves, sometimes out loud, intending for no other person to hear us, or mutter to limit the possibility that we might be caught at it. But just as important, the infant tendency to produce protophones is not primarily driven by attempts to communicate a particular emotional state (or anything else) to anyone. The great majority of protophones appear to be directed to no one (Long et al., 2020), but seem instead to constitute a kind of exploratory activity, where the infant investigates the nature of the vocal capacity itself and of the types of sounds that can be produced. Even infants born deaf produce massive numbers of protophones, with no evidence that the rate is lower across the first year than in hearing infants (Iyer and Oller, 2008). The conclusion seems inevitable that this vocal activity is pleasurable to infants, pursued in much the same way infants explore objects with their hands, eyes, and mouths, in an apparent attempt to understand the physical world. It is as if the human vocal capacity has come to be engaged for the purpose of playful exploratory activity, similarly to how the hands are engaged with the world in all primates. The vocal exploration yields an understanding of the acoustic properties resulting from infants' own vocal actions and the relations between those sounds and their kinesthetic accompaniments. This vocal exploratory/seeking behavior seems to be inherently reinforced just as other forms of play or Seeking behavior (see below) are deemed to be inherently pleasurable (Panksepp, 1982; Bekoff and Byers, 1998; Panksepp and Biven, 2012).

Of course there are other animals that produce abundant communicative vocalization. But something critically important for language appears to be absent in vocal activities of non-human apes: There appears to be no tendency to produce vocalizations exploratorily, playfully, seemingly for the sake of the sound experience itself, rather than for the sake of immediate communicative goals (Oller et al., 2019b). Another aspect of this apparent difference is that every human protophone type (by definition, the protophones do not include vegetative sounds or early infant cry or laughter) is produced with VFF, free to be expressed in any state of emotion or intent, whereas non-human vocalizations appear to be much more restricted to being produced as specific (although sometimes mixed) emotional expressions that primarily serve particular functional ends in the here and now.

Also in accord with the principle of VFF in adult humans, no immediate communicative intent (i.e., pursuit of a receiver reaction in the moment) is necessary for any particular language event type to occur, although clearly language would not have evolved had communication with others (both for immediate and long-term effects) not driven the selection of language abilities. We face an apparent paradox. Language is motivated and sustained by communication, but its nature requires that it be possible to use it "non-communicatively" – i.e., playfully and/or exploratorily. If it were not so, the capability would not be truly functionally flexible. So to form a foundation for vocal language, it is necessary for nature to select for a tendency to vocalize without any apparent immediate communicative purpose. Yet that tendency must have significant positive consequences for vocalizers in their own lifetimes. The selection advantage, we propose here depends on caregivers who notice the exploratory sounds of their infants, whether consciously or not, and who use the evidence of wellness inherent in those infant sounds (and the ones that are socially-directed as well) to modulate their investment in the infants' nurturance.

Empirical tests of the hypothesis that fitness signaling drives protophone production in modern infants can be envisioned in both behavioral and physiological domains. In the behavioral domain, one might predict significant correlations between rate of protophone production across individual infants (perhaps especially the rate of production of protophones when infants are comfortable) and level of parental investment in individual infant welfare. The correlations, we imagine might be most discernible in societies with high infant mortality. Low infant mortality in modern societies appears to have made it possible for many parents to invest most heavily in their least fit infants, in the hopes that all their offspring will be successful – so research in the most informative settings may be difficult to implement. In the physiological domain, one might predict increases in caregiver care-related neurochemicals such as oxytocin when they listen to protophone production. We are planning and encouraging research in both these domains.

We have argued that VFF is a foundation upon which all other aspects of vocal language depend (Oller et al., 2016). The argument is simple and intuitive, relying on the idea that some capabilities are required to develop early in order for others to develop later, because the later ones logically and practically depend on the earlier ones. The argument is supported empirically by the fact that human infants developing language actually go through the steps characterized in the natural logic. The first step in vocal language, as witnessed in longitudinal research, is the exercise of vocalization, copiously, playfully, and with no necessary expressed intent to communicate with others in the short term. This step seems obligatorily to involve VFF, since longitudinal research shows that endogenous, exploratory vocalization is always accompanied by VFF. In addition, without available endogenous infant vocalizations, caregivers would find no raw material with which to engage their infants in vocal interaction (Stern et al., 1975; Jaffe et al., 2001; Gratier et al., 2015), and consequently could not entrain infants in vocal turn-taking (Dominguez et al., 2016). Without vocal turn-taking, infants would not learn to participate in and contribute to protoconversation (Gratier and Devouche, 2011; Yoo et al., 2018). Without infant active participation in protoconversation, using vocalizations with VFF, systematic vocal imitation of new forms would not be possible (Jones, 2009; Long et al., 2019). Without these kinds of foundations, words and sentences could never be developed. This line of reasoning, illustrating that endogenous functionally flexible vocalization forms the initial platform for other critical developments necessary for language, has been presented in detail in other publications cited above, and is consistent with the well-documented facts of infant vocal and early language development summarized with citations in Oller et al. (2016).

Selection pressure on vocal flexibility must have affected hominins much more than closely related species because the functionally flexible capacity and inclination contrasts sharply with vocal inclinations in other apes. So, we are faced with the question: what was different about the situation where hominin vocal capacities must have passed through a phase transition into massively flexible vocal actions, while other apes remained more vocally constrained? The answer, we propose requires us to begin by taking stock of the nature of vocal communication in apes as well as in other non-human primates.

EMOTIONAL EXPRESSION AND VOCAL COMMUNICATION IN HUMANS AND OTHER PRIMATES

Our current view of the vocal systems of other primates is largely consistent with the original formulation of Darwin (1872), who proposed that vocal actions in many species, including apes and other primates are primarily emotional expressions. These expressions are sometimes complex and are clearly adaptable to circumstances (Snowdon et al., 1997; Crockford and Boesch, 2003; Hopkins et al., 2011), but they are fundamentally emotional nonetheless (Oller et al., 2019b).

The perspective on the role of emotion in communication has been informed recently by the work of Jaak Panksepp, who proposed seven basic emotions in mammals (Panksepp, 2011; Panksepp and Biven, 2012). Panksepp's perspective is discussed in detail in a separate paper in this volume (Griebel and Oller). Below, capitalized emotion terms are drawn from Panksepp's seven: Rage, Fear, Lust, Care, Panic (Isolation/Social need), Play (specifically Social Play), and Seeking. The most important point to emphasize here is that one of the seven, the Seeking system, is portrayed as a foundational emotion by which mammals (and presumably other metazoans) are driven to explore their worlds and are inherently rewarded by a sense of pleasure in the exploration itself. Note that the Seeking system can inspire exploratory, playful interaction with conspecifics and thus can activate the Social Play system, but most playful human infant vocalization seems to be independent of sociality, and thus, we propose that protophone production is primarily driven by the Seeking system rather than the Social Play system. A Seeking system is typically not present in other models of emotion (Eckman, 1994), but, we deem it a major advance in our understanding of emotion and of the basis for the massively endogenous and exploratory nature of human vocalization.

Vocalization in primates (except humans) is not explored for its own sake as far as we know, and thus it appears to be dissociated from the Seeking system in non-human primates. Instead, each vocalization type in mature non-human primates tends to be an expression of some other emotional state, selected to serve immediate, here-and-now functions. For example, some vocalization types tend to occur abruptly in response to Fear (distress and alarm calls), some to Rage (threats), some to Panic/Social Need (isolation calls, contact calls, and positive arousal calls), and some to Social Play (laughter). All these vocalization types can occur in circumstances as different as eating, traveling, and grooming, because all the emotional states can occur in any physical circumstance; e.g., in a feeding circumstance, competition for food can elicit the Rage system (possibly yielding vocal threats), the need to calm competitive tendencies regarding food can elicit the Care system (possibly yielding positive arousal/affiliation calls), and or perception of a predator can elicit Fear and/or Rage (possibly yielding a distress/alarm or threat call or a combination of them). We know of no evidence that any vocalization type in primates is confined tightly to any particular circumstance – rather emotions are inspired flexibly by events both internal and external to the organism, and their expression at each point in time may reflect the state of the producer more directly than the state of the environment. Importantly, emotional signals are flexible enough that they can sometimes be inhibited even when the corresponding triggering circumstances occur (Laporte and Zuberbühler, 2010; Owren et al., 2011).

So-called "predator-specific alarm calls" have been acknowledged, even in the earliest publications on the topic, to occur both in the circumstance of perceiving a predator and in intra-specific aggression (Seyfarth et al., 1980), and this point has been reconfirmed and elaborated in more recent revisiting of data regarding the species (the vervet monkey) on which the original alarm call research was done (Price, 2013; Price et al., 2015). Clearly the emotions of Fear and Rage are adaptable to eliciting vocal actions in widely different circumstances. Of course, in this argument, we do not dispute the idea that the physical environment can under some conditions elicit a particular emotion or a corresponding vocalization fairly reliably.

Vocalization in primates sometimes occurs in circumstances of low arousal, and in such cases one might ask if there is any emotion at all involved. Are such vocalizations equivalents to the protophones of human infants, displaying VFF? The answer must of course be determined empirically, and a trustworthy answer will depend on judgments of the functions of vocalizations occurring in their varying contexts. Our own research with three bonobo infants and their mothers in the first year (Oller et al., 2019b) suggests that some of the low arousal bonobo infant sounds, we observed (having occurred less than 1/10 as often as human protophones) were acoustically similar to some protophones, but we saw no evidence of human-like VFF. Essentially all the low arousal bonobo sounds that were produced and could be judged for function appeared to have negative valence (the infant trying to get back to mother and away from a harassing other bonobo, the infant whimpering for help after having climbed up on the cage and seemingly feeling unsure how to get down, and so on). These vocalizations could perhaps be attributed to Fear and/or Panic/Isolation. The judgment of valence in our research was based on how the infant acted before, during, and immediately after the vocalization, other events occurring at the time, and how the mother responded, often by picking the infant up and comforting or feeding him. Notably the bonobo mothers, while being very responsive physically, comforting infants or getting them out of trouble, never in 1,700 min of coded observation, responded to an infant vocalization with a vocalization of their own. Cases of bonobo infant vocalizations judged to have positive valence were deemed to be laughter, and not protophone-like (a laughter event in human infants is not treated as a protophone either). Perhaps most important, there was never a case of a vocalization at any intensity produced by a bonobo infant that was judged to be exploratory or playful – for criteria used in our human infant research to judge exploratory vocalization, see Long et al. (2020). In contrast, human infant protophones are abundantly judged to be exploratory, because they frequently show no sign of being directed to any one, are not judged to be based on discomfort, are not seen to have elicited immediate assistance, and are often produced when infants are alone in a room. At the same time, all the protophones of human infants show VFF and thus do occur on other occasions with social directivity, with signs of discomfort, with signs of delight, or in circumstances that elicit attention (often vocal attention) and/or help.

So far, there has been no convincing demonstration of functional flexibility in vocalizations of non-human primates, although there have been many demonstrations of *contextual* flexibility, that is, demonstrations that the same kind of sound occurs in different physical situations (de Waal, 1982; Harcourt et al., 1993; Biben and Bernhards, 1995; Bermejo and Omedes, 1999; Crockford and Boesch, 2003; Hopkins et al., 2011; Taglialatela et al., 2012). That a particular vocal type can occur in multiple physical situations can, of course, simply imply that similar emotional states occur in different physical situations.

One direct attempt to demonstrate VFF in adult bonobo vocalizations (Clay et al., 2015) did not actually address the issue, for two reasons: First, the study claimed to show that a particular vocal type (the peep) occurred in three situations: during aggression, traveling, and feeding. The authors interpreted the peeps as being negatively valenced during aggression, neutrally valenced during traveling, and positively valenced during feeding. This contextual variability does not, however, actually determine the function or emotional valence of the peeps occurring in these three different contexts. The same emotion that produces a peep could occur during any of the three contexts, in which case the function could be thought of, e.g., as an expression of annoyance (mild Rage) in all three cases or as an expression of Panic/Social Need in all three cases. It is untenable to assume that there exist one-to-one mappings of contexts to functions of vocalizations in primates (as was done in Clay et al., 2015) or of contexts to emotional states, since all emotions can occur in a variety of physical contexts, and correspondingly a variety functions can be served by vocal expression of those emotions in those varying contexts. This kind of flexibility is a defining characteristic of emotions in contrast to reflexes, which are more rigid and show shorter time frames from trigger to response. Emotions were evolved to allow flexible adaptations to important circumstances and challenges, and thus are subject to modification by learning and to cognitively-based adaptation (de Waal, 2019).

To prove VFF exists in a species, a workable approach is to demonstrate emotional valence variation from positive to negative in usage on different occasions of the same particular vocal type. Perhaps most important in order to demonstrate full VFF, it must be possible to demonstrate the occurrence of vocal events where there is *no discernible immediate function* – that is, the vocalization must be shown in some cases to be produced exploratorily and/or playfully. The peeps in Clay et al. (2015) were not shown to be produced exploratorily or playfully, and in fact no judgment was actually made about emotional valence (e.g., about facial expression, reaction of mother or other conspecifics, or other emotional indicators).

An additional problem with the study (Clay et al., 2015) was that it reported acoustic differentiation of the peeps occurring in the three contexts. If the data are correct, this acoustic demonstration undercuts the study's expressed goal, and the data did not demonstrate the existence of a single vocal type (a peep) with three functions, but three types of peeps, each with its own function. That humans might call all these sounds peeps does not prove they were all of the same vocal type to the bonobos, and the acoustic data suggest they could have indeed consisted of three different types to the bonobos.

Although there has been no convincing demonstration to our knowledge of VFF in non-human primates, the issue remains open to further investigation. We propose that for vocalization to become an object of exploration, it is necessary for natural selection to tie vocal capacities to an emotional system engendering actions that do not necessarily produce immediate benefits. If Panksepp was right, this would be the Seeking system, present in all mammals. Vocal inclinations in humans appear to have been evolved to be connected to the Seeking system in much the same way exploratory actions with the hands appear to have been connected to this emotional system in primates generally.

BARRIERS TO LANGUAGE EVOLUTION: THE PRESUMED ISSUE OF HONESTY

There must be barriers to language evolution, or we would not be the only creatures to have evolved it. The primary barrier that has been discussed in animal communication literature is based on the presumed competitive nature of signaling and its presumed resulting deceit. We are far from the first to express skepticism about this view or to outright reject it (see, e.g., Lachmann et al., 2001; Penn and Számadó, 2020). In Appendix A in the **Supplementary Material**, we address six key points that, in accord with our reasoning, counter the concerns and support the idea that the argument about deceit fails in providing an important barrier to either language evolution or stable communication in social-living non-humans. Here in the main text, we address what we believe to be the most fundamental reasons the idea of honest signaling as a barrier to language evolution is ill-conceived. These reasons are importantly related to the concept of VFF, as will be seen.

Consider the assumption that language is inherently honest. In fact, language is neither inherently honest nor inherently dishonest, a fact that can be illustrated with logical argument and examples alone. Acts of language are honest or dishonest depending on the circumstances they are intended to portray, and any mature speaker is capable of using language both ways. Perhaps the unsupportable claim that language needs to be honest is based on a confusion between the "meaning" of individual words, their semantics, and the way words are utilized to function (illocutionarily, see below) in communicative acts. The semantic meaning of a word, for example "rattlesnake," is dependent, not on truth or falsity, but on an understanding among speakers of English that the word refers to a particular class of animals. The word is neither honest nor dishonest in and of itself. The bond between the word and its semantic content is a convention sustained by speakers of a language over long periods (often centuries), not an individual assertion that might be falsified. But if an English speaker, who knows the difference between pythons and rattlesnakes, intentionally asserts that a particular python is a rattlesnake, the speaker is lying. It is not the word that is the lie, but the use of it to label an animal incorrectly. The same person might of course use the word truthfully and correctly on a different occasion. Importantly, in language we can also say things that are meaningful but are neither true nor false – and we do it very often. For example, suppose one says: "Please remove the rattlesnake." This could be a meaningful request; yet the request itself is neither true nor false. In writing the sentence about the rattlesnake, we have actually not made a request, but merely used a sentence as an example of a possible request. Nonetheless, the sentence, we have written uses meaningful English words in a meaningful and syntactically well-formed English sentence.

The fundamental misunderstanding that has been prevalent in animal communication literature based on the assumption that language is inherently honest can be unpacked and illuminated in the context of the Austinian distinction between illocutionary force and semantics (Austin, 1962). This distinction has been expanded in our own work so that it can apply not only to mature language, as it did for Austin, but also to human infant and animal communication (Oller, 2000; Griebel and Oller, 2008, 2014).

Illocutionary forces are the functions served in the here and now by communicative or potentially communicative acts. Illocutionary forces constitute the intentions that reflect underlying emotional/motivational states. Every production of a signal that has evolved to constitute a communication consists of at least one illocution, a performance of a communicative or potentially communicative act. For example, a scream emitted in Fear is an illocution, an "expression of Fear." Human infant cry can be portrayed illocutionarily as an "expression of distress." The hiss of a house cat can be viewed as a "threat." These illocutionary acts are not words; they possess no semantics and do not refer to anything, but instead express a state and/or a communicative intention. They are performances inspired by the state or intention in the present, and consequently they are neither true nor false.

On the other hand, any semantic (or symbolic) act consists of *both a semantic reference and at least one illocution*. If one says "rattlesnake," one may be performing a "labeling" illocution. Or with the same word, one might "correct" someone who had said "tree-root" (mistaking the snake for the root of a tree), and in so doing, one would produce two kinds of illocutionary functions in the same act, both a label and a correction. Saying "rattlesnake" could also be motivated by a fearful emotion, simultaneously invoking an "alarm" function along with the labeling function. Or one might say "rattlesnake" for the mere purpose of hearing the word, practicing it, or illustrating its pronunciation. Similarly, if one says "apple," one might intend merely a "labeling" of a fruit hanging from a tree. On a different occasion, one might use the same word to "request" that an apple be handed over, simultaneously labeling and requesting. With any word or phrase, we can perform many different illocutions. We can label, request, confirm, deny, alert, stipulate, mock, question (seek information), criticize, practice pronunciation, and so on.

But semantic acts always involve something in addition to illocution; semantics also includes the transmission of information encoded in the content of what is said. This semantic content is both transmitted in the here and now and in a broader sense is detached from the here and now. The word "rattlesnake" is a semantic entity that refers in English on every occasion of usage to a particular class of animals regardless of the intended illocution. The semantic content is independent of space and time, every time the word is produced. The semantic tie between a word and its conceptual content exists even in the absence of its being spoken. We can think a word or phrase and thus invoke the appropriate concept. The concept is invoked also regardless of the affective valence of the illocutionary act, that is, whether we produce the word or phrase with negative, positive, or neutral affect (for example, fearfully, delightedly, or exploratorily), differences that tend to correspond to different classes of illocutions.

We have contended that natural animal signals are limited to illocutionary functions and do not transmit semantic information (Oller and Griebel, 2015). We have thus far found no convincing contradictory evidence – only animals extensively taught by humans have been shown to transmit semantic content (Griebel et al., 2016). Thus, each naturally occurring animal communicative action is a performance (as far as we know), a mapping in the here and now, from an emotional or bodily state to the signal that expresses it as an illocution. The action does not "say" or "assert" anything, and thus can have no truth value.

For example, during mating season, a red deer who lowers his larynx and produces a sound involving lower resonances than if the larynx had been left in its rest state (Fitch, 2000) performs an illocutionary act we might call "advertisement" or "showing off." Did he lie by producing a sound with resonances suggesting a very large vocal tract and making himself sound larger than he is? No, because he did not say anything constituting an assertion that could be falsified. Male red deer in general produce mating calls with lowered larynges, and all of them do so for the same reason: the action makes them sound large and increases their probability of mating. Why? It appears female red deer choose to mate with males with deeper voices because deeper voices are related to greater body size and fitness. Any mutation that could have produced the inclination or capability to lower the larynx during mating calls could thus have been subject to runaway selection because it would have suggested great body size and fitness. The distinction between illocutionary force and semantics makes clear that illocutions are never true or false, because they are performances, not assertions. With semantic acts, however, we can indeed make claims about the world that may be subject to truth-value assessment.

Defenders of the idea that honest signaling is a barrier to language evolution might protest that they do not intend the term in animal communication theory to involve honesty as it can occur in language. Instead, they might argue that they only intend that the interaction between male and female deer involves the females being deceived into thinking the male deer they choose to mate with is bigger or more fit than he really is because he deceptively portrayed himself. This is an unnecessary conclusion. In the illocutionary interpretation, the male deer advertise by bellowing, and the female deer choose a mate on the basis of the effectiveness of the advertisement. Nothing can have been misinterpreted as true or false, because nothing was encoded semantically. Notice that all the advertising males lower their larynges. It is as if the honest signaling idea implies that all but one of them is lying. In that interpretation the females would have to be assumed to determine who is telling truth.

Propositions on the other hand (e.g., "there is a python" or "I am the biggest red deer in the forest") are semantic and can indeed involve "assertions" bearing semantic content, which can (at least in many circumstances) be determined empirically to be true, false, or ambiguous as to truth value. To produce such a proposition, one must invoke symbolic elements (typically sentences composed of words) to encode it. A human male can potentially try to impress a female by claiming with words to be rich and famous, which can be proven to be objectively true or false. The mating bellow of the male red deer, on the other hand, cannot be proven to be true or false.

Because every linguistic proposition is free to express a vast array of possible illocutionary forces, there is always a complex mapping possible between any linguistic symbols and their possible illocutionary functions. Many-to-many mappings also obtain between linguistic symbols and the different emotions that can be expressed by them, since the emotions motivate and supply flavoring for the illocutions. "Rattlesnake" can be produced contemptuously or admiringly. It can even be produced with flat affect for no purpose other than to speak the word. Or it can by produced to educate, teaching the label. The options are seemingly endless.

Deceit is of course possible through propositions – it logically *has* to be possible in language given the requirement of functional flexibility – and consequently deceit is among the possible illocutions of any proposition. This is not a weakness of language but an aspect of its power. If and only if a communicative system has the power to transmit both illocutions and semantic contents, can truth and falsehood be assessed. Language also makes it possible to create imaginary worlds, where talk about those worlds can involve only imaginary truths and falsehoods. Literary and cinematic fiction involves purely imaginary communications that can be evaluated for truth only in the context of the imagination. Did Star Trek's Captain Jean Luc Picard understand the Borg to be telling the truth, when it said "resistance is futile"? The question is not evaluable in the real world but is clearly meaningful and evaluable in the imaginary Star Trek world. The power of imagination supported by language yields vast possibilities in literature or cinema, but also in developing plans, providing explanations, coordinating actions, and so on.

In accord with our reasoning, the first step in selecting for such power, in moving beyond exclusively illocutionary communication, could not actually have involved selection for semantic capabilities. Rather a capability and inclination produced by selection had to form a foundation upon which a semantic system could later be built. This foundation, as we have argued above, involved the tendency in hominin infants to produce vocal fitness signals that had the (presumably unintentional) effect of revealing to caregivers their wellness and thus resulted in recurring nurturance of the infants through their long period of dependency. But crucially, the selection pressure was not on the quality of a *stereotyped* fitness signal, as in the case of the mating calls of the red deer, but on the tendency to explore various (not stereotyped) protophone types which could be interpreted as fitness signals. Selection for variety in fitness signaling can be found in mating and territorial calls in other species as well (e.g., birds and cetaceans), but to our knowledge these signals have never been shown to have VFF.

Importantly, one does not have to engage in fitness signaling intentionally in order for one's vocalizations to be interpreted as fitness signals. One wonders how many animals and humans produce fitness advertisements without even being aware of what they are doing. Does a male bird sing out of joy or because he intentionally wants to impress a female or a rival? Are humans always aware of the display functions that are served by things they say or how they say them? It would appear that selection pressure has created positive reinforcement (pleasure and joy) for singing, dancing, or whatever behavioral display is the advertisement proving to be effective in various animal species.

The evidence suggests that from the perspective of the human infant in the first months of life, most of protophone production is not an attempt to signal anything, but rather to engage in exploration or vocal play (as inspired by the

Seeking system), not directed to anyone. On other occasions, protophones do appear to be directed to a parent, in which case the vocalization might indeed represent an attempt on the infant part to bond with the parent – if not to signal fitness, at least to engage in social interaction in a playful way even in the first months of life (Gratier and Devouche, 2011; Dominguez et al., 2016; Yoo et al., 2018). From the caregiver perspective, infant vocalizations are presumably interpreted on many occasions in the same way the infant intends them, as explorations, for example. On other occasions protophones may be interpreted by caregivers as attempts to engage in social interaction. But at another level, the protophones heard by caregivers would seem always to supply fitness information, whether the infant intends them to supply such information or not. This result of protophone production is hypothesized in our approach to provide a basis for natural selection of infants ("parental selection" in the interpretation of Locke, 2006), who display their fitness through vocalizations, sometimes exploratory, sometimes interactive, sometimes emotionally expressive, but always in one way or another, providing rich information about infant state, well-being, and perhaps intelligence. To the extent that an infant vocalizes with the intention of "showing off," it might be appropriate to say the illocutionary force does indeed involve "fitness signaling." But judging infant intentions to this extent involves inferences that may be difficult to justify, just as it may be difficult to judge whether a bird intends his song to attract a mate or whether he merely intends to enjoy singing.

In summary, we see no barrier to the evolution of hominin vocal signaling because of an honesty issue. Language has to make both honest and dishonest communication possible, though a great many acts of language are not even evaluable with regard to honesty. The earliest communicative step away from the primate background in ancient hominis appears to have been the emergence of a capacity and an inclination to produce vocalizations as fitness signals long before there were words, and these fitness signals were neither true nor false.

POSSIBLE BARRIERS TO EVOLUTION OF VFF AND CIRCUMSTANCES THAT MAY HAVE HELPED OVERCOME THEM

What barriers would have actually inhibited natural selection of vocal exploration and VFF? One possibility is that there may be advantages to relative silence in order to avoid alerting predators or competitors. While, we know of no systematic investigation to tie it down, a silence pressure seems obvious. Perhaps for primates in general, the value of vocalizing freely was simply not high enough to get it off the ground in the face of a countervailing pressure for silence. Our hypothesis for over a decade (Oller and Griebel, 2005, 2006; Griebel and Oller, 2008), also advocated by Locke (2006, 2009), has been that hominin evolution occurred in circumstances where the value of vocalizing flexibly exceeded that of the pressure for silence.

In particular, we have proposed that the altricial hominin infant was in need of long-term care (Locke and Bogin, 2006), and thus came under especially intense pressure to provide fitness signals that could influence caregivers to provide long-term nurturance and protection. Altriciality is assumed in this reasoning to have been at least partly a product of bipedalism and the consequent narrowing of the birth canal, the "obstetrical dilemma" that is believed to have caused a necessary reduction of fetal brain-case size in hominins (Wells et al., 2012; Gruss and Schmitt, 2015). This reduction is assumed to have been accomplished by natural selection to slow development in hominins, resulting in smaller brains at birth and more altricial bipedal hominins than their quadrupedal cousins, who did not face the same obstetrical challenge (Bogin, 1999). Greater altriciality resulted, according to the reasoning, in greater need for long-term care along with greater advantages to fitness signaling by the altricial young.

Altriciality may not have provided the only selective pressure on flexible vocalization by infants. There are relatively few cooperative breeders among the primates, with humans and callitrichids (a New World group including marmosets and tamarins) being the only ones that are well-documented as showing both extensive care and provisioning by "alloparents" (Hrdy and Burkart, 2020). Interestingly both these groups are highly vocal, and the callitrichids show signs of greater flexibility of vocalization than other primate species (Snowdon and Cleveland, 1984; Snowdon and Elowson, 1999; Snowdon, 2004; Zuberbühler, 2011; Burkart et al., 2018), although the issue of possible VFF has not been directly evaluated in them. The callitrichids may be the only non-human primate group that babbles (Elowson et al., 1998). It has been argued that cooperative breeding is a setting that implies special pressure on infants to signal their needs and their fitness to a wide variety of possible caregivers, the alloparents. Increased volubility of these signals, especially when utilized in optimal circumstances, could surely enhance the prospects for such infants. We propose that relative altriciality and cooperative breeding may have co-evolved, with both supplying selective pressure on vocal fitness signaling in the hominin case.

One might object that a vastly new vocalization capability involving VFF is not the only way to supply fitness information. Fitness information is supplied by many features of an infant: skin color or texture, breathing pattern, frequency of crying, responsivity to touch or voice, and so on. Our argument is that the altricial hominin infant, especially in its cooperative breeding environment, was under more intense pressure to supply fitness information than other apes because of the longer developmental period ahead, a period during which there was absolute need for caregiver sustenance and the greater variety of caregivers. This enhanced pressure seems to have produced a new human feature, one where infants could supply fitness information to caregivers who were occupied with other tasks (see Falk, 2004 for an argument that language evolution was influenced by the common requirement of "putting the baby down" during foraging), and this pressure may have been redoubled in the circumstance of cooperative breeding because there were many potential caregivers. The protophones produced by hominin infants regardless of circumstances seem to have supplied a near constant source of well-being information, allowing hominin caregivers to assess the information and select individual infants for enhanced investment.

We assume the same pressures due to altriciality in the distant past exist also for modern humans and have suggested possible lines of empirical test (see above) of the idea that fitness signaling by protophones is noticed by caregivers, who respond both physiologically and behaviorally. Furthermore, a comparative evaluation of caregiving in species varying in altriciality could be informative. We predict that the more altricial the newborns of the species are at birth, the more intense the caregiving will be and especially the more attentive the caregivers will be to signals of fitness. On cooperative breeding, one might imagine correlational research focused on groups where degree of alloparenting differs. The prediction would be that protophone volubility will be positively correlated with the extent of alloparenting across groups. An additional prediction would be that both parents and alloparents would be sensitive to recognizing and responding with care to fitness signals.

Another factor that may have played a role in the hominin vocal inclination is suggested by research suggesting ancient hominin groups became larger than other ape groups in very distant time (Dunbar, 1993, 1996). With larger groups the premium on silence may have been mitigated somewhat by safety in numbers, allowing ancient hominins to be more subject to selection pressure on vocal fitness signaling and more social signaling in general. Dunbar's argument and that of Morris (1967) also emphasizes that as groups became larger, it became increasingly difficult to find enough time in the day to do all the grooming that primates seem to require to maintain peace in the group. Dunbar and Morris both proposed vocalization as having taken on a role similar to that of grooming in ancient hominins, because it was possible to vocally groom more efficiently, especially by including multiple recipients simultaneously.

An additional interpretation is that larger group sizes in hominins may have been in part made possible by emerging VFF. As vocalization became more frequent and more interactive, even if primarily between parents and infants, it surely would have been extended throughout the lifetime into utilization of vocalization to serve functions such as mating and alliance formation (vocal grooming). Such social vocalization usage may have fostered social cohesion with benefits not only to individuals but to groups of ancient hominins, whose numbers may have been able to expand in part because of the vocal connections and group commitments within their communities. In this interpretation, the vocal-grooming function may have co-evolved with the fitness-signaling function.

We are unaware of cross-species empirical research on this idea, but it could be tested for example, by evaluating the relative amounts of physical grooming and social vocalization in primate groups of varying sizes. Larger groups, other things being equal, might be expected to produce relatively larger amounts of social vocalization. Even in humans who live in hunter-gatherer societies, it may be possible to evaluate the relative amounts of physical grooming and social vocalization as a function of group size.

A critical feature of early hominin infant vocalization, selected as a fitness signal, according to our reasoning, was its connection with the Seeking system (see also Griebel and Oller, this volume), because it was this connection that motivated the copious production of protophones and gave them their flexibility. Vocalizations could be produced playfully without immediate utilitarian goals. The inclination to vocalize exploratorily appears to have been selected first and foremost as a form of investigation of the world, in this case the world of sound and its accompanying kinesthetics as produced by the vocal systems of the infants themselves. Assuming the vocal system is activated endogenously by the Seeking system, the activity can produce vocalizations varying in acoustic character for two reasons.

First, exploration can yield vocalizations that vary across a natural landscape of possible phonatory types corresponding to natural wells or "attractors" in a Waddingtonian landscape of vocal possibilities. Thus by self-organization, the exploration should produce variation and a tendency for categories to emerge. With increasing experience in exploration, the infant should learn to manipulate these categories, producing them repetitively and making them salient as categories. Indeed it has long been recognized that in modern human infants, several vocal categories tend to emerge in the first months: vowel-like sounds (vocants), squeals, growls, raspberries, and combinations of these (Zlatin-Laufer and Horii, 1977; Stark, 1978; Oller, 1981), and the repetition of each of these categories of sounds has been long recognized as a kind of vocal play, emerging at least by 5 months (Stark, 1980) but probably earlier (Jhang and Oller, 2017).

We have tested the identifiability of these protophone types auditorily (Oller et al., 2013) and based on human classification of spectrographic displays (Buder et al., 2008), and are currently comparing levels of agreement among human listeners compared with agreement between humans and automated acoustically-based identification. As a test of the extent to which exploration produces stable new sound types, we are currently involved in research on "clustering" of protophones of particular types (the tendency to produce particular types repetitively) across all-day recordings of typically developing infants and infants at risk for autism (Yoo et al., 2019b).

Second, acoustic properties characterizing particular emotional or affective states can modulate different protophone types so that each type, while maintaining acoustic signatures of its own, can simultaneously show acoustic variations tending to express differing emotional states. For example, a growl (which typically has harsh voice quality) might be produced on some occasions with no affective coloring but on other occasions with nasality and a whiny tone, along with a negative facial expression revealing discomfort. Intense discomfort might produce a louder version of the growl or an even more dysphonated and harsh version of the sound corresponding to a phonatory regime shift during at least part of the utterance. In a similar way, a vocant, which typically has normal phonation and reveals no affective positivity or negativity, might be produced with nasality and increased duration to signal distress (Yoo et al., 2019a). Squeals, which require high pitch, often in falsetto (or loft) register, can also be produced with affective neutrality, or can be colored by high intensity or the addition of intonational features suggesting distress.

We envision an evolutionary process where the tendency to vocalize flexibly in response to Seeking activation of the vocal system would have gradually become more frequent across (probably) millions of years. In successive generations, infants would have been increasingly fitness-signaling vocalizers, and the competition among infants for investment from caregivers would have persisted, a competition where fitness signaling vocalization would have always played a role (along with other fitness indicators, such as the appearance of skin health, coordination of movements, and so on). We also imagine that as the vocalizing infants grew up, they would have become more active users of vocalization in fitness signaling within their mature groups, with vocalization playing roles in mating and alliance formation.

Furthermore, as they became parents, the same individuals would have been sensitized to the value of vocalization as a signal of wellness, and they would have become increasingly attuned to noticing infant vocalizations as fitness signals. At some point, caregivers would have begun to elicit vocalizations by face-to-face vocal interaction, trying to gain access to information about fitness. Such face-to-face vocal interaction is common in human parents and infants (Brazelton et al., 1974; Cohn and Tronick, 1988; Jaffe et al., 2001; Hsu and Fogel, 2003), while never having been observed in other apes (Papoušek and Papoušek, 1983; Oller et al., 2019b). Not only would the endogenously-produced infant vocalizations have been useful indicators potentially benefiting infants who produced them, they would also have benefited their caregivers by providing a basis for allocating their investment energies. Thus fitness signaling was in the interest of both the infants and the caregivers. To the extent that there was competition, it was not primarily between caregiver and infant, but among infants who competed against each other for investment. A possible test of the competition among infants might be pursued in cases of multiple births. For example, one might seek to determine whether twins compete vocally in the sense that increases in protophone production by one twin produce increases in the other, independent of quotidian variations of production by each twin.

Another barrier to the evolution of vocal activity in the non-human primates seems likely attributable to relative lack of voluntary vocal control. The conclusion that voluntary vocalization is difficult for non-human primates has been noted in attempts to teach human-reared apes to produce anything resembling words – only the most minimal vocal "word learning" has been reported (e.g., Hayes and Hayes, 1951; Gardner et al., 1989). Similarly operant conditioning or social learning of vocalization in non-human primates has been shown to be difficult at best, with most authors emphasizing success in the realm of voiceless sounds, such as raspberries, smacking sounds, whistles, or whispered sounds (e.g., Marshall et al., 1999; Wich et al., 2009), with only minimal reported experience-driven modifications or modified uses of phonated vocalizations, and those modifications have applied to vocal types already existing in the relevant species repertoire (Sutton, 1979; Janik and Slater, 2000). The limits appear to be so severe that natural vocal learning in wild primates continues to be treated with a question mark about likely learning rather than an unambiguous positive conclusion (see, e.g., Crockford et al., 2004). A broad recent review concluded that the great bulk of vocal adjustments in non-human primates pertained to "vocal accommodation," involving adjustments to existing call structure based on environmental noise or conspecific vocalizations, without primates' learning to produce new sound types (Ruch et al., 2018).

In sharp contrast, a wide variety of other animals show clear vocal control and vocal learning, much more flexibly and easily achieved in the wild as well as in laboratory experiments. It appears that the vast majority of these animals either fly (e.g., songbirds, hummingbirds, parrots, and bats), have aquatic lifestyles (e.g., pinnipeds, dolphins, and whales), or have a history of aquatic or semi-aquatic lifestyles (e.g., elephants). In the case of the mammals that show impressive voluntary vocal control and vocal learning, most forage heavily or entirely in water, and must control their respiratory apparatus in such a way as to guage the amount of time necessary under water for each dive. This requirement imposes a necessity for voluntary control of the glottis (or other valve that manages respiratory flow). Since the glottis in primates is the apparatus that modulates respiration to create phonation, we propose that voluntary vocal control may have been facilitated by selection pressure on the voluntary control of the glottis that was naturally selected in hominins as a requirement of foraging by swimming and diving.

A hypothesis that ancient hominins lived at waterside, fishing and foraging in water by both wading and diving, and were heavily affected by selection pressures associated with these activities, appears to be gaining traction (Tobias, 2011; Attenborough, 2016). The idea has been on the table for many decades (Westenhoefer, 1942; Hardy, 1960; Morgan, 1997), but has been opposed by most of the community of paleoanthropology (e.g., Langdon, 1997), which is still primarily committed to the savannah hypothesis of human origins. Yet savannah living offers no integrated solution to explain the suite of characteristics that mark humans as remarkably distinct from their primate cousins, most importantly bipedalism, hairlessness, extensive subcutaneous fat, and voluntary phonatory control (for additional features and elaborations see, e.g., Niemitz, 2010; Gislen and Schagatay, 2011; Schagatay, 2011; Verhaegen et al., 2011). The idea is that ancient apes (perhaps the ancestors of both greater and lesser apes) spent significant periods of their evolution foraging in water, a pattern that may have influenced the evolution of preferential bipedalism on land (in hominins and gibbons) or other forms of special ambulation such as knuckle walking (in, e.g., chimpanzees and bonobos). In accord with reasoning by some supporters of this "waterside hypothesis," the hominins may have been the apes that stayed the longest in waterside living, experiencing to a much greater extent than other apes the special selection pressures of wading, swimming, and diving to forage. In accord with the hypothesis, hominins were evolved to be fully bipedal because wading places strong selection pressures on upright gait (see Kuliukas et al., 2009) and to possess the additional characteristics common in marine (and previously marine) mammals, as listed above.

This idea presents a possible basis upon which selection pressure on vocal fitness signaling could have taken hold in ancient hominin infants more easily than in other primates. If they were preadapted for more voluntary glottal control, they may have as a consequence been more susceptible to selection pressures on voluntary phonation than infants of other primate species. To our knowledge there has been no systematic research correlating the amount of diving done by various species with degree of vocal control. Such work should take into account the lifestyles of the species, since solitary creatures should not be expected to be as inclined to use social vocalization as gregarious species. Mating patterns should also be considered because mating songs also require vocal learning. Another test of the possible influence of hominin waterside living on vocal control could involve experimental studies of breath holding among primates. Instrumental conditioning research could conceivably make it possible to determine the degree to which apes and other primates can be taught to hold their breath.

The thoughts expressed here about possible impediments to the evolution of language are surely incomplete. Yet speculations and creative research on possible forces both favoring and inhibiting evolution of vocal flexibility hold promise in illuminating the origins of language.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

DKO and UG wrote the paper and conceived of the theoretical framework it expresses. Both authors contributed to the article and approved the submitted version.

FUNDING

The work reported here was inspired by a Focus Group funded by the Konrad Lorenz Institute for Evolution and Cognition (KLI) Research in Klosterneuburg, Austria and conducted at the KLI in the summer of 2019. The article was also supported by grants from the National Institutes of Health, R01 DC015108 and R01 DC011027 to DKO, and by the Plough Foundation Endowment which funds DKO's Chair of Excellence at the University of Memphis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.626138/full#supplementary-material

REFERENCES

Attenborough, D. (2016). The Waterside Ape. Retrieved from London, UK.

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

Bekoff, M., and Byers, J. A. (1998). *Animal play: Evolutionary, comparative and ecological perspectives*. UK: Cambridge.

Bermejo, M., and Omedes, A. (1999). Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatol.* 70, 328–357. doi: 10.1159/000021717

Biben, M., and Bernhards, D. (1995). "Vocal ontogeny of the squirrel monkey, *Saimiri boliviensis* peruviensis" in *Current topics in primate vocal communication*. eds. E. Zimmerman, J. D. Newman and U. Jürgens (New York: Plenum Press).

Bogin, B. (1999). *Patterns of human growth*. New York: Cambridge University Press.

Brazelton, T. B., Koslowski, B., and Main, M. (1974). "The origins of reciprocity: the early monther-infant interaction" in *The effect of the infant on its caregiver*. eds. M. Lewis and L. A. Rosenblum (New York: Wiley), 49–76.

Buder, E. H., Chorna, L., Oller, D. K., and Robinson, R. (2008). Vibratory regime classification of infant phonation. *J. Voice* 22, 553–564. doi: 10.1016/j.jvoice.2006.12.009

Burkart, J., Martins, E. G., Miss, F., and Zürcher, Y. I. S. (2018). From sharing food to sharing information: cooperative breeding and language evolution. *Inter. Stud. Soc. Behav. Commun. Biol. Art. Sys.* 19, 136–150. doi: 10.1075/is.17026.bur

Clay, Z., Archbold, J., and Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ* 3, e1124–e1120. doi: 10.7717/peerj.1124

Cohn, J. F., and Tronick, E. Z. (1988). Mother-infant face-to-face interaction: influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Dev. Psychol.* 24, 386–392. doi: 10.1037/0012-1649.24.3.386

Crockford, C., and Boesch, C. (2003). Context-specific calls in wild chimpanzees, *Pan troglodytes* verus: analysis of barks. *Anim. Behav.* 66, 115–125. doi: 10.1006/anbe.2003.2166

Crockford, C., Herbinger, I., Vigilant, L., and Boesh, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* 110, 221–243. doi: 10.1111/j.1439-0310.2004.00968.x

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

Dawkins, R., and Krebs, J. R. (1978). "Animal signals: information or manipulation?" in *Behavioural ecology: An evolutionary approach*. eds. J. R. Krebs and N. B. Davies (Sunderland: Sinauer Associates), 282–309.

de Waal, F. (1982). The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour* 106, 183–251.

de Waal, F. (2019). *Mama's last hug: Animal emotions and what they tell us about ourselves*. New York: Norton.

Dominguez, S., Devouche, E., Apter, G., and Gratier, M. (2016). The roots of turn-taking in the neonatal period. *Infant Child Dev.* 25, 240–255. doi: 10.1002/icd.1976

Dunbar, R. (1993). Coevolution of neocortical size, group size, and language in humans. *Behav. Brain Sci.* 16, 681–735. doi: 10.1017/S0140525X00032325

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

Eckman, P. (1994). "All emotions are basic" in *The nature of emotion*. eds. P. Eckman and R. J. Davidson (New York: Oxford University Press), 15–19.

Eilers, R. E., Oller, D. K., and Cobo-Lewis, A. B. (1997). Correlates of late onset canonical babbling in infants. *Emerg. Cogn. Lang.* 4, 7–10.

Eilers, R. E., Oller, D. K., Levine, S., Basinger, D., Lynch, M. P., and Urbano, R. (1993). The role of prematurity and socioeconomic status in the onset of canonical babbling in infants. *Infant Behav. Dev.* 16, 297–315. doi: 10.1016/0163-6383(93)80037-9

Elbers, L. (1982). Operating principles in repetitive babbling: a cognitive continuity approach. *Cognition* 12, 45–63. doi: 10.1016/0010-0277(82)90029-4

Elowson, A. M., Snowdon, C. T., and Lazaro-Perea, C. (1998). Babbling' and social context in infant monkeys: parallels to human infants. *Trends Cogn. Sci.* 2, 31–37. doi: 10.1016/S1364-6613(97)01115-7

Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *Behav. Brain Sci.* 27, 491–541. doi: 10.1017/s0140525x04000111

Fitch, W. T. (2000). The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. *Phonetica* 57, 205–218. doi: 10.1159/000028474

Fitch, W. T. (2004). "Evolving honest communication systems: kin selection and "mother tongues"" in *The evolution of communication systems: A comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 275–296.

Gardner, R. A., Gardner, B. T., and Drumm, P. (1989). "Voiced and signed responses of cross-fostered chimpanzees" in *Teaching sign language to chimpanzees*. *Vol. 165*. eds. R. A. Gardner, B. T Gardner and T. E. Van Cantfort (Albany, NY: State University of New York Press), 664–672.

Gislen, A., and Schagatay, E. (2011). "Superior underwater vision shows unexpected adaptability of the human eye" in *Was man more aquatic in the past? Fifty years after Alister Hardy-waterside hypotheses of human evolution*. *Vol. 1*. ed. M. Vaneechoutte (Sharjah, United Arab Emirates: Bentham Science), 164–172.

Gratier, M., and Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Dev. Psychol.* 47, 67–76. doi: 10.1037/a0020722

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Ebru, Y., and Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Front. Psychol.* 6:1167. doi: 10.3389/fpsyg.2015.01167

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

Griebel, U., and Oller, D. K. (2014). "Origins of language in a comparative perspective" in *Evolutionary science of human behavior: An interdisciplinary approach*. eds. P. J. Lafreniere and G. Weisfeld (Linus Learning: Rokokoma, NY), 257–280.

Griebel, U., Pepperberg, I. M., and Oller, D. K. (2016). Developmental plasticity and language: a comparative perspective. *Top. Cogn. Sci.* 8, 435–445. doi: 10.1111/tops.12200

Gruss, L. T., and Schmitt, D. (2015). The evolution of the human pelvis: changing adaptations to bipedalism, obstetrics and thermoregulation. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140063. doi: 10.1098/rstb.2014.0063

Harcourt, A. H., Stewart, K. J., and Hauser, M. (1993). Functions of wild gorilla 'close' calls. I. Repertoire, context, and interspecific comparison. *Behavior* 124, 89–122. doi: 10.1163/156853993X00524

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

Hayes, K. J., and Hayes, C. (1951). The intellectual development of a home-raised chimpanzee. *Proc. Am. Philos. Soc.* 95, 105–109.

Holmgren, K., Lindblom, B., Aurelius, G., Jalling, B., and Zetterstrom, R. (1986). "On the phonetics of infant vocalization" in *Precursors of early speech*. *Vol. 51*. eds. B. Lindblom and R. Zetterstrom, (New York: Stockton Press), 63.

Hopkins, W. D., Taglialatela, J., and Leavens, D. A. (2011). "Do apes have voluntary control of their vocalizations and facial expressions?" in *Primate communication and human language*. eds. A. Vilain, J. L. Schwartz, C. Abry and J. Vauclair (Amsterdam: John Benjamins Publishing), 206–226.

Hrdy, S. B., and Burkart, J. M. (2020). The emergence of emotionally modern humans: implications for language and learning. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 1–11. doi: 10.1098/rstb.2019.0499

Hsu, H. C., and Fogel, A. (2003). Social regulatory effects of infant non-distress vocalization on maternal behavior. *Dev. Psychol.* 39, 97–991. doi: 10.1037/0012-1649.39.6.976

Iyer, S. N., Denson, H., Lazar, N., and Oller, D. K. (2016). Volubility of the human infant: effects of parental interaction (or lack of it). *Clin. Linguist. Phon.* 30, 1–19. doi: 10.3109/02699206.2016.1147082

Iyer, S. N., and Oller, D. K. (2008). Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss. *Volta Rev.* 108, 115–138. doi: 10.17955/tvr.108.2.603

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., and Jasnow, M. D. (2001). *Rhythms of dialogue in infancy: Coordinated timing in development*. Chicago: University of Chicago Press, 66.

Janik, V. M., and Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Anim. Behav.* 60, 1–11. doi: 10.1006/anbe.2000.1410

Jhang, Y., and Oller, D. K. (2017). Emergence of functional flexibility in infant vocalizations of the first 3 months. *Front. Psychol.* 8:300. doi: 10.3389/fpsyg.2017.00300

Jones, S. S. (2009). The development of imitation in infancy. *Philos. Trans. R. Soc. B* 364, 2325–2335. doi: 10.1098/rstb.2009.0045

Koopmans-van Beinum, F. J., and van der Stelt, J. M. (1986). "Early stages in the development of speech movements" in *Precursors of early speech*. *Vol. 37*. eds. B. Lindblom and R. Zetterstrom (New York: Stockton Press), 50.

Kuliukas, A. V., Milne, N., and Fournier, P. (2009). The relative cost of bent-hip bent-knee walking is reduced in water. *Homo J. Comp. Hum. Biol.* 60, 479–488. doi: 10.1016/j.jchb.2009.09.002

Lachmann, M., Számadó, S., and Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proc. Natl. Acad. Sci.* 98, 13189–13194. doi: 10.1073/pnas.231216498

Langdon, J. (1997). Umbrella hypotheses and parsimony in human evolution: a critique of the aquatic ape hypothesis. *J. Hum. Evol.* 33, 479–494. doi: 10.1006/jhev.1997.0146

Laporte, M. N. C., and Zuberbühler, K. (2010). Vocal greeting behaviour in wild chimpanzee females. *Anim. Behav.* 80, 467–473. doi: 10.1016/j.anbehav.2010.06.005

Lee, C. -C., Jhang, Y., Relyea, G., and Oller, D. K. (2017). Subtlety of ambient-language effects in babbling: a study of English-and Chinese-learning infants at 8, 10, and 12 months. *Lang. Learn. Dev.* 13, 100–126. doi: 10.1080/15475441.2016.1180983

Locke, J. L. (2006). Parental selection of vocal behavior: crying, cooing, babbling, and the evolution of language. *Hum. Nat.* 17, 155–168. doi: 10.1007/s12110-006-1015-x

Locke, J. L. (2009). Evolutionary developmental linguistics: naturalization of the faculty of language. *Lang. Sci.* 31, 33–59. doi: 10.1016/j.langsci.2007.09.008

Locke, J. L., and Bogin, B. (2006). Language and life history: a new perspective on the evolution and development of linguistic communication. *Behav. Brain Sci.* 29, 259–325. doi: 10.1017/S0140525X0600906X

Long, H. L., Bowman, D. D., Yoo, H., Burkhardt-Reed, M. M., Bene, E. R., and Oller, D. K. (2020). Social and non-social functions of infant vocalizations. *PLoS One* 15:e0224956. doi: 10.1371/journal.pone.0224956

Long, H. L., Oller, D. K., and Bowman, D. D. (2019). Reliability of listener judgments of infant vocal imitation. *Front. Psychol.* 10:1340. doi: 10.3389/fpsyg.2019.01340

Marshall, A. J., Wrangham, R. W., and Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Anim. Behav.* 58, 825–830. doi: 10.1006/anbe.1999.1219

Masataka, N. (2001). Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate-limiting constraint on the emergence of canonical babbling. *Dev. Sci.* 4, 158–164. doi: 10.1111/1467-7687.00161

Mehl, M. R., Vazire, S., Ramírez-Esparza, N., Slatcher, R. B., and Pennebaker, J. W. (2007). Are women really more talkative than men? *Science* 317:82. doi: 10.1126/science.1139940

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

Morris, D. (1967). *The naked ape*. New York: Dell.

Niemitz, C. (2010). The evolution of the upright posture and gait—a review and a new synthesis. *Naturwissenschaften* 97, 241–263. doi: 10.1007/s00114-009-0637-3

Nyman, A., and Lohmander, A. (2018). Babbling in children with neurodevelopmental disability and validity of a simplified way of measuring canonical babbling ratio. *Clin. Linguist. Phon.* 32, 114–127. doi: 10.1080/02699206.2017.1320588

Oller, D. K. (1981). "Infant vocalizations: exploration and reflexivity" in *Language behavior in infancy and early childhood*. ed. R. E. Stark (New York, North Holland: Elsevier), 85–104.

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., and Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl. Acad. Sci.* 110, 6318–6632. doi: 10.1073/pnas.1300337110

Oller, D. K., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Lee, C. -C., et al. (2019a). Preterm and full term infant vocalization and the origin of language. *Sci. Rep.* 9:14734. doi: 10.1038/s41598-019-51352-0

Oller, D. K., and Eilers, R. E. (1982). Similarities of babbling in Spanish-and English-learning babies. *J. Child Lang.* 9, 565–578. doi: 10.1017/S0305000900004918

Oller, D. K., and Eilers, R. E. (1988). The role of audition in infant babbling. *Child Dev.* 59, 441–449. doi: 10.2307/1130323

Oller, D. K., Eilers, R. E., Basinger, D., Steffens, M. L., and Urbano, R. (1995). Extreme poverty and the development of precursors to the speech capacity. *First Lang.* 15, 167–188.

Oller, D. K., and Griebel, U. (2005). "Contextual freedom in human infant vocalization and the evolution of language" in *Evolutionary perspectives on human development*. eds. R. Burgess and K. MacDonald (Thousand Oaks, CA: Sage Publications), 135–166.

Oller, D. K., and Griebel, U. (2006). How the language capacity was naturally selected: altriciality and long immaturity. Commentary on Locke, J. and Bogin, B: language and life history: a new perspective on the evolution and development of linguistic communication. *Behav. Brain Sci.* 29, 293–294. doi: 10.1017/S0140525X06389066

Oller, D. K., and Griebel, U. (2015). Vocal development in the human infant and in our closest primate relatives. Paper presented at the Human Behavior and Evolution Society, Columbia, MO, University of MO.

Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., et al. (2019b). Language origin seen in spontaneous and interactive vocal rate of human and bonobo infants. *Front. Psychol.* 10:729. doi: 10.3389/fpsyg.2019.00729

Oller, D. K., Griebel, U., and Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language Topics in Cognitive Science (topiCS), Special Issue: New Frontiers in Language Evolution and Development. *Top. Cogn. Sci.* 8, 382–392. doi: 10.1111/tops.12198

Owren, M. J., Amoss, R. T., and Rendall, D. (2011). Two organizing principles of vocal production: implications for nonhuman and human primates. *Am. J. Primatol.* 73, 530–544. doi: 10.1002/ajp.20913

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behav. Brain Sci.* 5, 407–422. doi: 10.1017/S0140525X00012759

Panksepp, J. (2011). Toward a cross-species neuroscientific understanding of the affective mind: do animals have emotional feelings? *Am. J. Primatol.* 73, 545–561. doi: 10.1002/ajp.20929

Panksepp, J., and Biven, L. (2012). *The archaeology of mind: Neuroevolutionary origins of human emotion*. New York: W. W. Norton & Company.

Papaeliou, C., Minadakis, G., and Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *J. Speech Lang. Hear. Res.* 45, 311–317. doi: 10.1044/1092-4388(2002/024)

Papoušek, H., and Papoušek, M. (1983). Biological basis of social interactions: implications of research for an understanding of behavoiural deviance. *J. Child Psy.* 24, 117–129. doi: 10.1111/j.1469-7610.1983.tb00109.x

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., and Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: canonical babbling status and vocalization frequency. *J. Autism Dev. Disord.* 44, 2413–2428. doi: 10.1007/s10803-014-2047-4

Penn, D. J., and Számadó, S. (2020). The handicap principle: how an erroneous hypothesis became a scientific principle. *Biol. Rev.* 95, 267–290. doi: 10.1111/brv.12563

Price, T. (2013). Vocal Communication within the Genus Chlorocebus: Insights into Mechanisms of Call Production and Call Perception. (PhD), Georg-August-Universität, Göttingen.

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5:13220. doi: 10.1038/srep13220

Ruch, H., Zürcher, Y., and Burkart, J. M. (2018). The function and mechanism of vocal accommodation in humans and other primates. *Biol. Rev.* 93, 996–1013. doi: 10.1111/brv.12382

Schagatay, E. (2011). "Human breath-hold diving ability suggests a selective pressure for diving during human evolution" in *Was man more aquatic in the past? Fifty years after Alister Hardy-waterside hypotheses of human evolution. Vol. 1.* ed. M. Vaneechoutte (Sharjah, United Arab Emirates: Bentham Science), 120–147.

Scheiner, E., Hammerschmidt, K., Jürgens, U., and Zwirner, P. (2002). Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *J. Voice* 16, 509–529. doi: 10.1016/S0892-1997(02)00127-3

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28, 1070–1094. doi: 10.1016/S0003-3472(80)80097-2

Snowdon, C. T. (2004). "Social processes in the evolution of complex cognition and communication" in *Evolution of communication systems: A comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 131–150.

Snowdon, C. T., and Cleveland, J. (1984). "Conversations" among pygmy marmosets. *Am. J. Primatol.* 7, 15–20. doi: 10.1002/ajp.1350070104

Snowdon, C. T., and Elowson, A. M. (1999). Pygmy marmosets modify call structure when paired. *Ethology* 105, 893–908. doi: 10.1046/j.1439-0310.1999.00483.x

Snowdon, C. T., Elowson, A. M., and Rousch, R. S. (1997). "Social influences on vocal development in new world primates" in *Social influences on vocal development*. eds. C. T. Snowdon and M. Hausberger (New York, NY: Cambridge University Press), 234–248.

Stark, R. E. (1978). Features of infant sounds: the emergence of cooing. *J. Child Lang.* 5, 379–390. doi: 10.1017/S0305000900002051

Stark, R. E. (1980). "Stages of speech development in the first year of life" in *Child phonology. Vol. 1.* eds. G. Yeni-Komshian, J. Kavanagh and C. Ferguson (New York: Academic Press), 73–90.

Stark, R. E. (1981). Infant vocalization: a comprehensive view. *Infant Med. Health J.* 2, 118–128.

Stern, D. N., Jaffe, J., Beebe, B., and Bennett, S. L. (1975). Vocalizing in unison and in alternation: two modes of communication within the mother-infant dyad. *Ann. N. Y. Acad. Sci.* 263, 89–100. doi: 10.1111/j.1749-6632.1975.tb41574.x

Stoel-Gammon, C. (1992). "Prelinguistic vocal development: measurement and prediction" in *Phonological development: Models, research, implications*. eds. C. Ferguson, L. Menn and C. Stoel-Gammon (Parkton, MD: York Press, Inc.), 439–456.

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

Taglialatela, J. P., Reamer, L., Schapiro, S. J., and Hopkins, W. D. (2012). Social learning of a communicative signal in captive chimpanzees. *Biol. Lett.* 8, 498–501. doi: 10.1098/rsbl.2012.0113

Tobias, P. V. (2011). "Revisiting water and hominin evolution" in *Was man more aquatic in the past? Fifty years after Alister Hardy-waterside hypotheses of human evolution. Vol. 1.* ed. M. Vaneechoutte (Sharjah, United Arab Emirates: Bentham Science), 3–15.

Verhaegen, M., Munro, S., Puech, P. -F., and Vaneechoutte, M. (2011). "Early hominoids: orthograde aquarboreals in flooded forests?" in *Was man more aquatic in the past? fifty years after alister hardy - waterside hypotheses of human evolution. Vol. 1*. ed. M. Vaneechoutte (Sharjah, United Arab Emirates: Bentham Science), 67–81.

Vinter, S. (1994). L'analyse du babillage: une contribution au diagnostic de surdité? *Approche Neuropsychologique des Apprentissages chez l'Enfant* 6, 232–238.

Wells, J. C. K., DeSilva, J. M., and Stock, J. T. (2012). The obstetric dilemma: an ancient game of Russian roulette, or a variable dilemma sensitive to ecology? *Am. J. Phys. Anthropol.* 149, 40–71. doi: 10.1002/ajpa.22160

Westenhoefer, M. (1942). *Der Eigenweg des Menschen*. Berlin: Mannstaedt & Co.

Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., and Shumaker, R. W. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates* 50, 56–64. doi: 10.1007/s10329-008-0117-y

Yoo, H., Bowman, D., and Oller, D. K. (2018). The origin of protoconversation: an examination of caregiver responses to cry and speech-like vocalizations. *Front. Psychol.* 9:1510. doi: 10.3389/fpsyg.2018.01510

Yoo, H., Buder, E. H., Bowman, D. D., Bidelman, G. M., and Oller, D. K. (2019a). Acoustic correlates and adult perceptions of distress in infant speech-like vocalizations and cries. *Front. Psychol.* 10:1154. doi: 10.3389/fpsyg.2019.01154

Yoo, H., Ha, S., Lee, C. -C., Ramsay, G., and Oller, D. K. (2019b). A cross-cultural analysis of clustering of infant vocalizations across the first year of life: Results combined across the Memphis/Emory studies and Hallym University in S. Korea. Paper presented at the American Speech-Language-Hearing Association, Orlando, FL.

Zlatin-Laufer, M. A., and Horii, Y. (1977). Fundamental frequency characteristics of infant non-distress vocalization during the first 24 weeks. *J. Child Lang.* 4, 171–184. doi: 10.1017/S0305000900001586

Zuberbühler, K. (2011). "Cooperative breeding and the evolution of vocal flexibility" in *Oxford handbook of language evolution*. eds. K. R. Gibson and M. Tallerman (New York, NY: Oxford University Press), 71–81.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 19 March 2021
doi: 10.3389/fpsyg.2021.602635

Why Are No Animal Communication Systems Simple Languages?

Michael D. Beecher[1,2]*

[1] *Department of Psychology, University of Washington, Seattle, WA, United States,* [2] *Department of Biology, University of Washington, Seattle, WA, United States*

Individuals of some animal species have been taught simple versions of human language despite their natural communication systems failing to rise to the level of a simple language. How is it, then, that some animals can master a version of language, yet none of them deploy this capacity in their own communication system? I first examine the key design features that are often used to evaluate language-like properties of natural animal communication systems. I then consider one candidate animal system, bird song, because it has several of the key design features or their precursors, including social learning and cultural transmission of their vocal signals. I conclude that although bird song communication is nuanced and complex, and has the acoustic potential for productivity, it is not productive – it cannot be used to say many different things. Finally, I discuss the debate over whether animal communication should be viewed as a cooperative information transmission process, as we typically view human language, or as a competitive process where signaler and receiver vie for control. The debate points to a necessary condition for the evolution of a simple language that has generally been overlooked: the degree of to which the interests of the signaler and receiver align. While strong cognitive and signal production mechanisms are necessary pre-adaptations for a simple language, they are not sufficient. Also necessary is the existence of identical or near-identical interests of signaler and receiver and a socio-ecology that requires high-level cooperation across a range of contexts. In the case of our hominid ancestors, these contexts included hunting, gathering, child care and, perhaps, warfare. I argue that the key condition for the evolution of human language was the extreme interdependency that existed among unrelated individuals in the hunter-gatherer societies of our hominid ancestors. This extreme interdependency produced multiple prosocial adaptations for effective intragroup cooperation, which in partnership with advanced cognitive abilities, set the stage for the evolution of language.

Keywords: animal communication, language evolution, animal cognition, animal language studies, information

Edited by:
Irene M. Pepperberg,
Harvard University, United States

Reviewed by:
Carel ten Cate,
Leiden University, Netherlands
Slawomir Wacewicz,
Nicolaus Copernicus University in Toruń, Poland
Erich David Jarvis,
Duke University, United States

****Correspondence:***
Michael D. Beecher
beecher@uw.edu

Specialty section:
This article was submitted to Evolutionary Psychology, a section of the journal Frontiers in Psychology

Received: *03 September 2020*
Accepted: *18 February 2021*
Published: *19 March 2021*

Citation:
Beecher MD (2021) Why Are No Animal Communication Systems Simple Languages? Front. Psychol. 12:602635. doi: 10.3389/fpsyg.2021.602635

INTRODUCTION

Research programs on animal communication systems in nature have proceeded essentially independently of research programs endeavoring to teach language to animals. This is surprising in light of the early, well-known efforts to relate these two research streams, especially by Hockett (1960) and Marler (1961). These efforts spurred two questions. First, can animals be taught human language, even a simplified version? Second, do the natural communication systems of any animals rise to the level of simple language? Research since then has indicated that these two questions may have different answers: I would suggest a provisional yes to the first, and a provisional no to the second. If this view is correct, it raises a further question: why, then, if some animals can master a version of language, don't they use this capacity in their natural communication system? In this paper I address this paradox, and make some suggestions toward its resolution.

My paper is divided into four parts. First I consider the main "design features" of language proposed by Hockett as a basis for evaluating language-like properties of animal communication systems. Hockett concluded that some animal communication systems have some of these design features, but none of them have all the design features, especially the key ones. I will designate an animal communication system as a 'simple language' system using a variation on the definition of Hewes (1973): "language [is] any system of animal communication which exhibits most of the design features set forth by Hockett" (Hewes, 1973, p. 5). I narrow this definition by identifying four design features – semanticity, arbitrariness, learnability and cultural transmission, and productivity – as necessary for the system to be classified as a simple language. Second, I discuss bird song, a case where several but not all of the key design features are present. I will focus on one specific case of a song-based communication system that is clearly complex and nuanced, but nevertheless lacks three key design features, semanticity, arbitrariness and productivity. Third, I consider the debate, not yet fully concluded, over whether animal communication should be conceived of as a process of information transfer or as manipulation of receiver by the signaler. The debate is germane to our more specific question because it provides a clue as to why we find no simple languages among animals despite the apparent capacity for it in at least some of them. Finally, I suggest that although there appear to be at least some animals with the cognitive capacity for a language-like communication system, none of them have a social system with extreme interdependency among individuals on the scale of that which existed in the hominid hunter-gatherer system. I argue that this extreme interdependency was a necessary condition for the evolution of human language.

DESIGN FEATURES OF LANGUAGE

In this section I consider the extent to which the most important design features of human language are found in animal communication systems. I use Hockett's (1960) design features as a basis for comparison of natural animal communication systems with human language. Although Hockett's design features may have limited use as a theoretical framework for modern evolutionary linguistics (Wacewicz and Żywiczyński, 2015), it is a useful starting point for the comparative analysis of this paper. I have winnowed Hockett's original design features down to the few I consider the most fundamental ones that can be used to directly compare human language with animal communication systems.

Specialization: The Purpose of Linguistic Signals Is Communication and Not Some Other Biological Function

Specialization, in Hockett's sense, is the first defining feature of a communication system, no matter how simple or complex it might be. Otte (1974) defines communication signals as traits "fashioned or maintained by natural selection because they convey information to other organisms"(Otte, 1974, p. 385). I discuss the vigorous debate over the 'information' aspect of this definition in Section "Communication: Information or Influence? Mutual Benefit or Manipulation?", but debaters on both sides would agree that this definition captures the key difference between true communication signals on the one hand, and tactical behaviors or inadvertent cues on the other. For example, while we might describe an individual delivering a blow to a potential opponent as 'sending a message,' we mean this only in a metaphorical sense. This behavior is primarily tactical, that is, the individual delivering the blow will directly benefit it if its opponent responds by backing down. If instead of delivering a blow the individual had said "I'm going to kill you," or growled, or barked, or hissed, we would recognize these as true communication signals, having been shaped by natural selection for the purpose of (literally) sending a message, and requiring adaptations in the receiver as well.

Hockett listed prevarication – the ability to transmit misinformation, i.e., to lie or deceive – as one of his many design features, albeit a minor one, a corollary almost. In Section "Communication: Information or Influence? Mutual Benefit or Manipulation?", I will argue that we should consider prevarication to be a fundamental, indeed foundational feature of animal communication systems: communication in animals is shaped by the tension between the sender's and receiver's interests, and truth in communication is not a given, but rather, when it occurs, hard won.

Semanticity: Specific Signals Are Directly Tied to Certain Meanings

To say that a communication system is *semantic* is to say that it uses signals to represent particular things or actions. A well-known example in animals are alarm signals given in response to different predators. We can say in such cases that each of these signals represents one of several different predators, or more precisely, the appearance on the scene of one of these predators. For example, vervet monkeys have three different alarm calls for three different classes of predators: raptors, terrestrial mammals and snakes, predators which depend on an element of surprise to capture the monkey. In response to an aerial predator, such as a martial eagle, a monkey emits 'cough' calls and sender and receivers take shelter in dense bushes or near the core of a tree. In response to leopards, a monkey emits a 'bark' call and the monkeys climb up to the tip of tree branches where leopards cannot safely go. Finally, if a monkey spots a dangerous snake, such as a python, it emits a 'chutter' call and the group gathers around the snake, standing upright and harassing it until it leaves the area. Although the vervets use these same signals in other contexts (e.g., intergroup fights) to represent different things, the modification of signal meaning in different contexts occurs in human language as well, and does not negatively impact the representational quality of these signals (Seyfarth et al., 1990; Price et al., 2015). Indeed, it is not unusual for an animal to use a particular signal to mean different things in different contexts (Smith, 1997), similar to some words meaning totally different things within different sentences.

Nevertheless, I will argue later in this paper that the semanticity of animal communication systems is limited: although some things are represented by animal signals, the number of things is generally small. Attempts to catalog the number of different things signaled in animal communication systems typically top out at 25 or so (vervet monkeys, Struhsaker, 1967; Japanese macaques, Green, 1975; review in Hauser, 2000). The limitation does not appear to be due to production constraints (the ability to produce enough distinct signals or to recombine enough of them to enlarge the signal set) or to perceptual-cognitive constraints.

Arbitrariness: Languages Are Made Up of Arbitrary Symbols Which Have No Intrinsic or Logical Connection to What They Represent

A distinctive feature of human language is that not only are words semantic, they are arbitrarily so. We could equally well call dogs ‘cats’ and cats ‘dogs,’ or any other two words, so long as sender and receiver knew the convention, a point illustrated by the existence of the many different languages of the world. These signals seem totally arbitrary with respect to what they signify, and in theory they could be interchanged without problems, so long as senders and receivers were both aware of the convention. How about animal signals? It appears that in theory we could interchange the vervet alarm signals without problems, provided of course that the receivers were aware of the ‘convention’ (i.e., were hard-wired appropriately). Identity signals – indicating species or individual identity, and occasionally group or kinship – are perhaps the most common animals signals that unequivocally have the arbitrariness feature.

But many, perhaps most, animal signals are not arbitrary. Signals used in agonistic and mate attraction contexts are typically “more of” signals, i.e., more effective signals are louder, longer, bigger, brighter, flashier, designed to impress or to shock and awe. I am unaware of any clear example where the reverse is true, where the more effective signal is the one that is less conspicuous, for example, a softer sound, a more subdued color, a less vigorous display. An apparent exception might be the ‘quiet song’ sung by many songbirds in intense conflict situations, but this typically happens only when the bird is close to its opponent so that the quiet song is audible to the receiver (Searcy et al., 2014); ‘normal’ song is loud because it is a long-distance signal. Moreover, quiet song is typically different in other respects besides loudness, for example, having some elements seen only in quiet song, such as very high frequency elements.

Other animal signals are simple extensions or slight modifications of tactical behaviors, e.g., of attack behavior in agonistic situations. For example, a threat signal in many mammals is the open mouth display, where the teeth, the canines notably, are prominently displayed. Ethologists called this a ‘ritualized’ display (Lorenz, 1966), i.e., one that has been modified by natural selection to be a display, since the mouth is held open, and attack withheld, rather than being the beginning of an actual attack. Another common threat signal is the raising of the hair or feathers, making the animal appear larger. Again, while these actions are plausibly considered ritualized displays, they are not arbitrary signals. If they were, you would also find cases where animals threaten by closing their mouths, or by making themselves appear small. In short, animal signals functioning to impress an opponent or potential mating partner are usually inherently impressive, not arbitrarily selected to represent threat or desirability. Any naïve observer viewing a ritualized dominance interaction between two wolves (or dogs) would have no difficulty determining which animal was dominant and which was subordinate. An upright animal, with its hair raised, its tail raised, and staring at its opponent inherently appears dominant, whereas one with a flattened, slinking body, hair down, tail down, and looking away from the opponent, inherently appears subordinate.

Many epigamic signals – signals designed to attract a mate and induce her to mate – are bright, striking ornaments, often ones that function like supernormal stimuli (e.g., the tail of the long-tailed widowbird, Andersson, 1982). Many epigamic signals are energetically expensive and highly skilled behaviors, such as the complex male courtship dances of wolf spiders and jumping spiders (Hebets and Uetz, 1999; Elias et al., 2012). The motor performance revealed in these sorts of displays likely reflect whole-organism performance relating to survival, and thus should be good indicators of individual signaler quality. There is considerable evidence that females choose mates in nature based upon their evaluations of male motor performance (reviewed in Byers et al., 2010). The relevant point here is that these signals are not arbitrary, but inherently reflect the trait signaled: signaler quality.

Even in the example par excellence of communication of information about the external world – the honeybee dance language – the signals are not quite so arbitrary as generally assumed. For example, if the dance is done outside the hive, where the sun is visible, the bee dances with respect to the actual position of the sun, rather than with respect to the vertical (Gould, 1975). That is, outside the hive, the symbology is not truly arbitrary. Moreover, the distance to the target is represented by the duration of the straight run – the further the distance, the longer the run – so this is at least partially non-arbitrary as well.

Although the words in human language are arbitrary – the existence of different languages is the clearest evidence on this point – they may be expressed in such a way to amplify or otherwise modify their meaning, as for example a loudly shouted “no” indicating stronger conviction. But what would be considered an extra-linguistic feature for humans is often the primary message in animals. For example, the initial stage of a battle between two male red deer consists of a roaring contest (Clutton-Brock and Albon, 1979). This vocal signaling duel does far more than simply establish that each animal is a male conspecific ready to defend or fight for the harem – this undoubtedly was perceived by both parties before the contest began – rather, how loud and how long an individual roars establishes how motivated and formidable he is, and is used by the receiver to decide whether to continue the fight or depart. Similarly, the plumage ornaments and courtship dance of a male golden-collared manakin do far more than simply identify species and sex – that is simply the necessary first step – the brightness of the ornament and the skill of the dance determine whether the receiver, the female, will choose to mate with this particular male or continue her search for the best possible mate (Stein and Uy, 2006; Barske et al., 2011).

In summary, although we have examples of animal signals that are totally arbitrary, many others – perhaps most? – are not. I would add that to date we have found nothing comparable to the many different human languages, which are a consequence of the arbitrariness feature. We do find geographical dialects in animals (e.g., Marler and Tamura, 1964; Wright and Dahlin, 2018), but as the name implies, these are relatively minor variations on the basic signal set, nothing like the wholesale variation seen in human languages.

Learnability and Cultural Transmission

Human language is both learned and taught. Most animal communication systems are neither. A well-known exception to this generalization are the learned vocal communication signals of several taxa, most notably the oscine passerines (songbirds), hummingbirds and parrots among birds, and cetaceans and at least some bat species among mammals (reviews in Janik, 2014; Knornschild, 2014; Nowicki and Searcy, 2014). Evidence for vocal learning and cultural transmission in some other birds and mammals as well (Walcott et al., 2006; Kroodsma et al., 2013; Stoeger and Manger, 2014; Garland and McGregor, 2020; Barker et al., 2021) suggests that this ability may lie closer to the surface than is generally assumed, but at least at the present time, vocal learning is thought to be rare in animals. Later in this paper I return to the best-studied example of vocal learning, song learning in songbirds.

Where the communication signals are learned, we should expect to find dialects, geographical variation in the signals. The occurrence of dialects is one criterion for identifying the occurrence of learning and potentially evidence for the arbitrariness design feature. An example that may illustrate the arbitrary nature of dialects is the recently-discovered modification of the song in eastern white-throated sparrows to resemble the typical song of western white-throated sparrows. Investigators have traced this change to eastern birds learning the western version of the song on the migration grounds, where individuals of the two populations mix (Otter et al., 2020). Most eastern birds now sing the 'western' version of the song on the breeding grounds, illustrating that the details of the song structure are not crucial for its function. Although Otter et al. (2020) suggest that this change might have been driven by a preference on the part of eastern females, they give no evidence for this hypothesis, nor plausible basis for it.

Perhaps even rarer in animal communication systems than learning is teaching. The commonly accepted criteria for demonstrating teaching in non-human animals are that (1) teachers should modify their behavior in the presence of the learner, (2) this change in behavior should result in no immediate benefit to the teacher, and (3) the learner should acquire a behavior quicker or better as a result (Caro and Hauser, 1992). In song-learning studies the birds from whom the young bird learns its song are conventionally referred to as 'tutors,' and although live birds are invariably more effective song tutors than recorded song (review in Beecher, 2017), the term 'tutor' is used purely as matter of convenience. In fact, in the most common context for song learning in nature, young birds learn from older birds who are or will be their territorial rivals, a very different context from language learning in young humans, where 'tutors' are typically relatives or other interested parties who ultimately (but not immediately) benefit from tutoring. Nevertheless, even in the common songbird case where the young bird learns from territorial rivals, bird song tutoring would fit all three criteria for teaching if in fact the older bird reduces his usual aggression when a young bird appears on his territory, increases his counter-singing with the young bird in such a way as to facilitate learning, and benefits down the road from this tutoring (for example, the two cooperate in mutual defense of their territories, or against predators, or refrain from extra-pair mating with one another's mates). We have indirect evidence for song learning/teaching in song sparrows: mutual survival is greater in young birds and their primary tutor-neighbor (the one from whom they learn most of their songs) the more songs the two of them ultimately share, i.e., the more songs the tutee learned from the tutor, or the tutor taught the tutee (Beecher et al., 2020).

Productivity: By Combining a Small Number of Meaningless Units Into Larger Meaningful Signals, a Sender Is Capable of Producing Meaningful Statements About Virtually Anything

The sense in which I am using this term is captured by Hauser (2000, p. 448): "the power of [human] language comes from our capacity to take meaningless syllables and combine them into an unbounded number of meaningful words, and then take these words and combine them into an unbounded number of meaningful expressions (Chomsky, 1986; Studdert-Kennedy, 1998)." I will define productivity as recombining a smaller number of basic signal units to produce a larger number of signals, and thus, messages. Indeed, semanticity (representation) and productivity are probably the two central features of human language: by combining basic phonetic units into larger meaningful units, and combining these units further via syntactical rules, we can say almost anything.

Animal communication systems are not productive in this sense, and this is the primary reason we do not refer to them as languages. We would be impressed if a vervet could say something like "Grab your infant and run from the leopard coming from the west but watch out for the python who likes to hide in the bushes just to the east of you." A human can say this kind of thing easily, combining a relatively small number of atomic units (phonemes) into very large number of basic signals (words) and combining these into a very large set of possible communications. I note that while there is some controversy in phonetics about exactly what are the units of productive combination, there is agreement that all natural languages (including sign language) are made up of meaningless atomic units that are combined into larger meaningful wholes (Zuidema and de Boer, 2009).

Instead of productivity, we could describe the communication system in terms of information capacity. The information capacity of human language is essentially infinite, in the sense that, in theory, we can communicate virtually anything. Our motor, sensory and cognitive capacities obviously will reduce how much information actually gets transmitted and received. But still, the fact is that we can transmit an enormous amount of information with language. Attempts to measure information capacity or information transmission in animals, on the other hand, have given rather modest results. Two estimates of the information about distance and direction in the honeybee dance language have given a high value of 14.9 bits (Gould, 1975) and a low value of 7.4 bits (Schürch and Ratnieks, 2015). My group has estimated the information capacity of the call signature system that parents of the colonial cliff swallow use to find their offspring in their large breeding colonies (Medvin et al., 1993). We estimated the capacity as 8.76 bits, and the estimate would be somewhat larger if we included information that can be derived from visual differences among cliff swallow chicks (Stoddard and Beecher, 1983). The information capacity of human language of course is orders of magnitude larger than this.

We certainly find the potential for productivity in bird song. For example, most songbirds have multiple songs (song 'repertoires'), and the different songs are made up of different syllables or notes in different orders, and these smaller units can be used in more than one song. Still, although the units are there, and although songbirds may possess the cognitive capacity to comprehend hierarchical structuring in vocal signals (Gentner et al., 2006; but see van Heijningen et al., 2009), they do not use these capacities to form different songs *representing* different things. As Hauser (2000, p. 450) puts it, "in contrast to the recombination of words into sentences by humans, the output of songbird recombination does not change its meaning." A minor exception are some songbirds who use some song types in a territorial defense context and others in a mate attraction context (e.g., Byers, 1996). As discussed in the next section, theories on the function of song repertoires abound, but they all agree that the different songs function simply to provide diversity, rather than to represent different things.

Summing Up

Table 1 summarizes the conclusions of this section. The natural communication systems of animals fall short of human language on a number of the key design features of language. They come closest on semanticity, where signals sometimes represent things in the external world or within the signaler, and the signals are sometimes truly arbitrary. However, more commonly animal signals are not arbitrary but inherently meaningful, e.g., an animal making itself appear large is more frightening than an animal making itself appear small. Most animal communication signals and responses are neither learned nor culturally transmitted. And, so far as we know, no animal communication has the *sine qua non* of language: productivity.

BIRD SONG: COMPLEXITY WITHOUT PRODUCTIVITY

The oscine passerines (songbirds) are one of the rare animal taxa in which individuals learn their vocal communication signals. In most animals, these vocal signals are 'hard-wired,' that is, they develop normally whether or not the animal is exposed to them early in life. It has long been noted that vocal learning in songbirds has many similarities to language learning in humans (Marler, 1970; Doupe and Kuhl, 1999). These similarities include the following. (1) The young bird needs to be exposed to normal species vocal signals in order to produce them as an adult. (2) The sensory phase of song learning precedes the motor phase. (3) Auditory feedback (which can be abolished by deafening) is

TABLE 1 | Key design features of communication systems (after Hockett, 1960, pruned and combined).

Found in animals?	Design feature	Comment
Yes	Specialization. The purpose of linguistic signals is communication and not some other biological function.	True of animal communication systems, but this is essentially by definition.
Yes but limited	Semanticity. Specific signals are directly tied to certain meanings.	Clear example are the alarm calls given to different classes of predators in a number of species. But the number of different things signaled is typically very small.
Yes but rare	Arbitrariness. There is an arbitrary relationship between a signal and its meaning. There is no inherent relationship between the form of a signal and what it refers to.	Animal signals are sometimes arbitrary. Often they have inherent meaning that can be readily perceived by a naïve observer, e.g., signals used in mate attraction or agonistic encounters that are designed to impress or shock and awe.
Yes but rare	Learnability and Cultural transmission.	Human language is learnable, teachable and culturally transmitted. Bird song appears to be one of the few animal examples that passes at least two of these criteria (teaching still not established).
No	Productivity (based on Arbitrariness, Discreteness and Duality of patterning): language made up of small meaningless units which can be combined into many larger meaningful units which can be combined to say virtually anything.	Some animals appear to have the motor and cognitive capacity for a productive, language-like communication system but they do not use this capacity to develop language-like communication systems.

necessary for the translation of memorized sensory input into motor production. (4) Vocal learning is most efficient in (and sometimes restricted to) a sensitive period early in life. (5) There are specialized parts of the brain dedicated to the vocal control system. (6) Song is socially learned and culturally transmitted, and in at least some cases it may be actively taught (e.g., Carouso-Peck and Goldstein, 2019; Beecher et al., 2020). While notable differences exist among songbird species with regard to the normal progression of song learning (Beecher and Brenowitz, 2005), these six features are essentially true for all of the many songbirds that have been studied to date.

Despite the notable parallels between bird song learning and human language learning, none of the many studies endeavoring to teach a version of human language to animals have focused on songbirds. This is all the more surprising given the language learning shown by Alex the African Gray Parrot, a member of another avian taxon with vocal learning, the psittacines (Pepperberg, 1981, 1987). Moreover, songbirds have strong cognitive capacities, a highly-developed vocal production mechanism, and a vocabulary of basic sound units in their song that rivals or exceeds the basic sound units of human language. There are even songbird species that can mimic human speech sounds (e.g., Hill Mynah birds). On the face of it, all the requisites would seem to be there to support a simple language in a songbird.

What Is the Function of a Song Repertoire?

In contrast to well-studied white-crowned sparrows and zebra finches, in most songbird species an individual bird will sing multiple songs (has a song 'repertoire'). For example, song sparrows typically have nine (plus or minus two or so) very different songs. Each of these songs is made up of 5 or 6 distinct elements, and the order of these elements is important (Horning et al., 1993). The songs do not have individual signatures and the nine or so songs in a song sparrow's repertoire are as different among themselves as would be a collection of songs taken at random one from each of nine or so different birds (Beecher et al., 1994). Song sparrows are somewhere on the middle of the song repertoire complexity scale: many species have larger and even more complex song repertoires. The key point for this discussion is that song repertoires provide clear potential for productivity, as song sparrows and many other songbirds have as many or more distinct units in their vocal communication systems (e.g., about 100 in indigo buntings, Thompson, 1970; and in swamp sparrows, Marler and Pickert, 1984) as there are in human language (a typical language has 40–45 phonemes).

The most popular hypothesis about song repertoires for north temperate zone songbirds – where only males sing – is that they are an epigamic signal produced by males to attract females and that larger repertoires are more attractive than smaller ones (Catchpole, 1987; Searcy and Yasukawa, 1996; MacDougall-Shackleton, 1997; Collins, 2004). Focusing on just the well-studied song sparrow, the evidence for this hypothesis is mixed (Searcy, 1984; Reid et al., 2004; Hill C. E. et al., 2011). The handicap principle, discussed in the next section, would suggest that if large song repertoires are preferred, it is because they are an indicator of some aspect of male quality. Reid et al. (2005) found support for this idea: song repertoire size in male song sparrows correlated with enhanced cell-mediated immune response (CMI) and relative heterozygosity. Anderson et al. (2017) hypothesized that female song sparrows might prefer large-repertoire males because this feature is an indicator the overall learning ability of the male. However, they found no correlations between repertoire size (or two other measures of song learning ability) with an overall measure of learning ability (based on five different learning tasks). I should note, however, that a correlation of vocal learning ability with both overall learning ability and mating success has been found in another songbird, the Satin Bowerbird, a vocal mimic: in this case the vocal learning ability is the ability of males to mimic the calls of other local bird species, both the number of species mimicked, and the accuracy of the mimicry (Coleman et al., 2007; Keagy et al., 2009).

According to another hypothesis, song repertoires play a role in territorial competition, which in north temperate zone songbirds, where only males sing, is largely male-male competition, but outside the north temperate zone where both sexes sing, is pair-pair competition (e.g., Levin, 1996; Langmore, 1998; Logue and Gammon, 2004). There are several hypotheses as to how repertoires might work in the territorial competition context. Song is used by most territorial songbirds at least in part as a keep-out signal, to 'post' their territory. Kroodsma (1988) argues that the vocal diversity provided by a repertoire functions to hold the attention of territorial competitors by dishabituating them to the territory owner's singing, i.e., by holding their attention. As one piece of evidence, he points to a positive correlation between repertoire size and population density in marsh wren populations, and also to the finding that birds in denser populations cycle through their songs faster, again a behavior that should reduce habituation (Kroodsma, 1977). In contrast, song sparrows sing their much smaller repertoires with eventual variety, i.e., singing each one of their song types many times before switching to another type, and this would seem to argue against the dishabituation hypothesis. In western, resident populations of song sparrows, song repertoires may function primarily to provide a bird with songs matching all (or most) of his neighbors, and thus potential individualized replies to each one of them (Beecher et al., 1997; and see next section).

Although as this brief discussion indicates, the theoretical debate has not yet concluded, the take-away point is that none of these hypotheses view song repertoires as a form of semantic communication. Rather they view repertoires as having a direct effect on the receiver (dishabituation), or as permitting individualized replies to multiple neighbors, or as quantitative signals with inherent rather than semantic meaning, that is, more songs (or more song syllables) are simply more effective.

I should add that most single-song species appear to have the potential to develop song repertoires yet do not tap into this potential. For example, when examined over an entire population, indigo buntings have a repertoire of over a 100 distinct song syllables, yet a given individual uses just 6–8 of these in the single song it develops (Rice and Thompson, 1968; Thompson, 1969; Baker and Boylan, 1995).

An Example: Communication in a Negotiation Context

Although the different songs in a bird's repertoire do not have different meanings, a bird having a song repertoire can still use the different songs to communicate in more subtle, nuanced ways than might at first be suspected. In this section I describe one such case: how song sparrows use the songs in their song repertoire to negotiate territorial disputes. The general point I will make is that their communication system is surprisingly complex and versatile, despite being neither semantic nor productive. Although I will not attempt to generalize to all songbirds given the incredible diversity of the song communication systems seen in this group (Beecher and Brenowitz, 2005), I suspect that this conclusion – complexity without productivity – applies broadly to songbirds, and perhaps to all animals.

Song sparrows have a territorial system like that found in many animals and typical of many songbirds. An individual carves out a territory where the mated pair will nest and raise their young, doing most of their feeding on the territory. Suitable habitat is typically densely occupied by conspecifics, so territorial disputes can arise during both the establishment and maintenance stages. The relationship between territorial neighbors can become relatively non-hostile once established, however, on the principle that the enemy you know is better than the enemy you don't know, generally referred to as the 'Dear Enemy' relationship (Fisher, 1954; Akçay et al., 2009, 2010; Beecher and Akçay, 2014). Because in territorial animals, neighbors have no fences, neighbors need to renegotiate territory boundaries from time to time. Negotiation can progress into fighting but avoiding fighting may benefit both parties and this common interest favors reliable signaling. Therefore, as I will discuss in Section "Communication: Information or Influence? Mutual Benefit or Manipulation?", we should expect to find some degree of honest communication concerning not only fighting ability (resource-holding potential) but also motivation to fight (e.g., at a particular point in time, one party may have more to lose than the other).

Song sparrows in western, resident populations use their repertoires in a complex way to carry out territory negotiations. Although they will engage in serious fights, established neighbors use their signaling system to avoid fighting if possible. Before fighting they typically give their high-level threat signals, wing waves and soft song (Searcy and Beecher, 2009; Searcy et al., 2014; Akçay et al., 2015a). But before reaching this stage, they use the songs in their repertoires to escalate or de-escalate the dispute following a set of 'conventions' predicated on which songs the two birds happen to share (Beecher et al., 1996, 2000; Burt et al., 2001, 2002; Beecher and Campbell, 2005; Akçay et al., 2011; Templeton et al., 2012; Akçay et al., 2013, 2015b). Because western song sparrows learn songs from their neighbors in the area to which they disperse after fledging, a bird typically shares some of his songs with each of his immediate neighbors. The set of songs he shares with one neighbor is typically different from the set he shares with another. A partial example is shown in **Figure 1**. For example, if we represent the different songs of a bird with different capital letters, and the shared songs of neighbors with the same capital letter, then Bird 1 might share his song types A, B, and C with his neighbor Bird 2, his song types C, D, and E with another neighbor, his song types E and F with a third neighbor, and finally G, H, and I with no neighbors (e.g., the bird he learned these songs from may have died). A typical territorial negotiation might occur as follows. Suppose Bird 1's mate finds an ideal place to build her nest just over the previously-established boundary with Bird 2. Bird 1, aiming to establish this new boundary, moves to that point and sings at his neighbor. Typically the two birds would still be a considerable distance apart at this point and out of sight of one another (territories are large and song is a long-distance signal). Although Bird 1 could sing any one of his 9 songs to Bird 2, in this circumstance he would typically 'address' Bird 2 by singing one of their shared types, A, B, or C. Let us say bird 1 sings B. Bird 2 can escalate by replying with his B' (i.e., his most similar song to Bird 1's B). This 'type match' is a low-level threat signal and would be the first step in escalation. Alternatively, he could 'confirm' without escalating by replying with A' or C' ('repertoire matches', Beecher et al., 1996). Note that this type of reply is only possible if Bird 2 knows Bird 1 well enough to know which songs they share and which songs they don't. Finally, rather than type-matching or repertoire-matching, Bird 2 can de-escalate by singing one of his unshared types, e.g., D, E, F, G, H or I. Singing an unshared type is better than not singing at all because it signals that although the singer is not engaging, he is on territory and has heard his neighbor; it is a signal likely used for example when the bird is busy feeding recently-fledged young. If Bird 2 does type match bird 1 (sings B'), Bird 1 in turn can continue to sing that song type ('stay on type'), or he can de-escalate by switching to another shared song (A or C, 'repertoire match'), or de-escalate further by switching to an unshared type (e.g., D or E), or disengage totally by stopping singing.

Each 'convention' – type matching, repertoire matching, staying on type, switching to an unshared type – has a distinct signaling function in this graded signaling system, with both type matching and staying on type when type-matched signaling a readiness to escalate, repertoire matching signaling recognition of the sender and engagement but stopping short of escalation, and switching to an unshared type signaling de-escalation. The system while not in itself resolving anything, does give the neighbors time to defuse the situation or work out a compromise. Note, however, that the semantic content is limited. No particular song in the repertoire means a particular thing. A song's meaning is defined entirely by the context of who the receiver is, and even then there are essentially only three meanings, roughly 'back off,' 'I hear you and know who you are,' and 'I'm busy now.'

Summing Up

Songbirds check several of the design feature boxes and they would appear to have the potential to use their songs in a productive way, i.e., to use their signaling system to say many things. However, despite considerable debate concerning the function of song repertoires, the different repertoire hypotheses all agree on one point: that the function of the vocal diversity

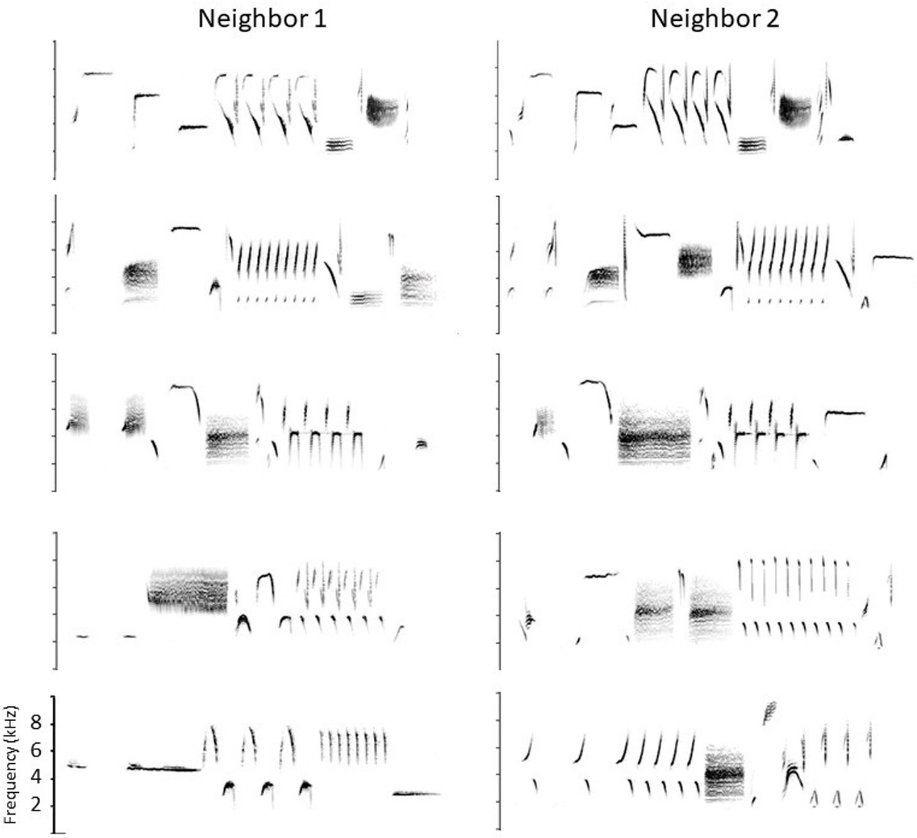


FIGURE 1 | Partial song repertoires of two neighboring birds. Shared songs are shown in the top three rows, and four of their unshared songs in the bottom two rows (they are arbitrarily paired). Frequency scale: 0–10 kHz. Songs are 2–3 s long.

is diversity *per se*, not the transmission of different messages with different songs. Perhaps even more surprising, many single-song species have large song syllable repertoires an individual could tap into, but instead each individual uses just several of these syllables to develop its single song. No songbird rearranges its multiple song syllables into different songs that signal different things. I echo here the conclusion of Fitch and Jarvis (2013, p. 502): although songbirds (and parrots) have vocal learning and a complex vocal repertoire, they do not "use their songs to communicate combinatorial propositional meanings, i.e., semantics.". Songbirds may use their repertoires in subtle, nuanced ways, as with the song sparrow hierarchical signaling system I described above, but what the system achieves seems better described as the management of behavioral conflict than as an impressive transmission of information. That is, the system may function well, but it does not function like a language.

COMMUNICATION: INFORMATION OR INFLUENCE? MUTUAL BENEFIT OR MANIPULATION?

In this section I discuss the debate within the field about the fundamental nature of animal communication. I believe this debate has provided us with a key to understanding why we find no examples of a simple language among the many communication systems of non-human animals, and true language only in the human animal.

We can trace the real beginning of the field of animal communication to the classical ethologists (e.g., Tinbergen, 1952). The ethologists provided detailed descriptions of animal signaling systems in nature, developed theories about the underlying proximate causes (e.g., sign stimuli, innate release mechanisms, and fixed action patterns) and evolutionary processes (e.g., ritualization), and most relevant here, established the view of animal communication as – like human language – an information transfer process. On the question of the function of animal signaling systems, they took a group-selectionist perspective: the benefit that a signaling system provided went not to signaler or receiver *per se*, but to the species (see Tinbergen, 1964 definition in **Table 2**).

Following the revolution of the 1960's and 1970's first known as sociobiology (Wilson, 1975) and subsequently as behavioral ecology (Krebs and Davies, 1978), natural selection came to be viewed as acting on individuals, rather than species or groups (Williams, 1966). For some researchers, the shift from naïve group selection to individual selection did not entail a significant change in view: it was simply assumed that signaler and receiver

TABLE 2 | Definitions.

Tinbergen, 1964	"One party… emits a signal, while the other party… responds in such a way that the welfare of the species is promoted."
Marler, 1968	In "true communication… both parties seek to maximize the efficiency of information transfer."
Otte, 1974, p. 385	Signals: "behavioral, physiological, or morphological characteristics fashioned or maintained by natural selection because they convey information to other organisms"
Dawkins and Krebs, 1978, p. 283	"Communication is said to occur when an animal, the actor, does something which appears to be the result of selection to influence the sense organs of another animal, the reactor, so that the reactor's behavior changes to the advantage of the actor."
Green and Marler, 1979, p. 73	"Communication consists of the transmission of information from one animal to another."
Krebs and Dawkins, 1984, p. 401	They call the sender role the 'manipulator' and the receiver role the 'mind-reader.' "The manipulator role is selected to alter the behavior of others to its advantage, the mind-reader role to anticipate the future behavior of others."
Smith, 1997, p. 11	Communication: "any sharing of information between entities—in social animals, between individual animals"
Bradbury and Vehrencamp, 1998, p. 3	True communication: "information exchange from which both sender and receiver benefit."
Maynard Smith and Harper, 2003, p. 3	A signal is "any act or structure that alters the behavior of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved."
Owren et al., 2010, p. 771	Animal Signaling: "the use of specialized, species-typical morphology or behavior to influence the current or future behavior of another individual."

both benefited from the transmission of information, and so this basic parallel with human language was maintained (see **Table 2** definitions of Marler, 1968; Otte, 1974). The assumption of mutual benefit seemed natural in cases where sender and receiver have a strong common interest, e.g., the honeybee 'dance language' where scout and recruit are both working toward the same end, to provide food for their relatives in the hive. But as investigators began considering the many cases where signaler and receiver have conflicting interests, such as in agonistic encounters over an indivisible resource, they began to question the mutual-benefit, information transmission view. They asked two questions about such cases. First, do both parties have to benefit? Second, do we need to even talk about 'information transmission'? Isn't the signaler simply selected to manipulate (or influence) the behavior of the receiver to its advantage? The manipulation viewpoint was famously developed by Dawkins and Krebs (1978) who argued that rather than expecting signalers to signal honestly, we should expect them to manipulate the receiver to their own advantage, e.g., to convince opponents to retreat, or potential partners to mate with them.

Since the Dawkins and Krebs (1978) paper, the debate has continued as to whether it is justified or productive to conceptualize animal signaling as an information transmission process in which both parties benefit. Simplifying somewhat, I will distinguish between the Information Transmission and Manipulation approaches to animal communication. Strong arguments on the manipulation side since Dawkins and Krebs (1978) include Krebs and Dawkins (1984), Owings and Morton (1998), Scott-Phillips (2008), Rendall et al. (2009), and Owren et al. (2010). Strong arguments on the information side over this same period include Green and Marler (1979), Smith (1997), Bradbury and Vehrencamp (1998), Searcy and Nowicki (2005), Carazo and Font (2010), Seyfarth et al. (2010), and Wiley (2013). Definitions from some of these sources are included in **Table 2**.

In conceiving of signaling as manipulation, Dawkins and Krebs (1978) essentially treated the communication interaction like a zero-sum game. This seems reasonable in cases like disputes over an indivisible resource (a food item, a territory, and a mate), and also in epigamic selection, where a male tries to persuade a female to mate with him now rather than to continue searching for a possibly better male. Although the manipulation view was enlightening in many respects, as originally presented it had a serious weakness: it gave no agency to the receiver. While it was sensible to expect signalers to signal for their own benefit, why should we expect receivers to be passive in these evolutionary scenarios, especially if being manipulated by the signaler is costly? Rather, we should expect receivers to show 'sales resistance' to signals that carry misinformation or are pure propaganda ("I am the best," "I will fight you to death"). Indeed, receivers can do more than simply ignore signals that do not benefit them: they can require signals that do benefit them, even if those signals are costly to the sender. For example, in many species males must sing or call to attract a female for mating. If the male does not vocalize, potential female receivers will simply not engage. Moreover, these vocal signals may attract predators, a cost borne by the signaler but not the receiver. Indeed, the most effective or most-preferred signals may be the most costly, e.g., most conspicuous not just to the intended receiver but to predators as well. This is the case for a male túngara frog (Ryan and Rand, 1990). Males attract females to mate with a 'whine' call or a 'whine-chuck' call. When a male adds chucks to his calls, he not only attracts more females, but also predators: frog-eating bats that home in specifically on the chucks. Similarly, a calling male field cricket attracts more females than does a silent male, but he also attracts more parasitoid flies, and louder calls attract both more females and more parasitoid flies (Cade, 1975). In some populations the rate of fly parasitism is so high that males have lost the ability to sing (Zuk et al., 2006). As another example, territorial animals often vocalize as a "keep-out" signal. When a territorial songbird is deprived of its voice, however, potential rivals show up and proceed to take over its territory (e.g., McDonald, 1989).

If we reframe our view of the communication system as beginning with the implicit requirement that the receiver imposes on the signaler—to signal—rather than with the signal itself, it is apparent that receivers can be conceived of as manipulating signalers, and in the 'receiver manipulation' view, the potential costs to the sender are secondary to the potential benefits to the receiver. A possible benefit for the female túngara frog – the receiver in our example – might be a shorter search time in navigating to the male who adds the more localizable chucks to his calls, perhaps lessening her vulnerability to predation.

The receiver manipulation view prompts us to consider how the receiver might demand a more honest signal. There are two related possibilities. First, the receiver can selectively attend to signals that are inherently honest due to physical constraints. For example, in many frogs and toads, size is the most important weapon in male battles over mating opportunities and size is reliably predicted by the pitch of the animal's vocalization: larger animals give lower-pitched calls. Davies and Halliday (1978) showed that playback of low-pitched calls was sufficient to discourage smaller males from entering into battle with an apparently larger male. A second way to require a more reliable signal has generally been discussed under the rubric of the 'handicap' principle. This principle was first proposed by Zahavi (1975), modified and formalized by Grafen (1990), given the intuitively pleasing graphical formulation by Johnstone (1997) shown in **Figure 2**, and is still being subjected to further modification and clarification (e.g., Penn and Számadó, 2018). But the basic principle is straight-forward, and can be verbalized as follows: signals whose degree of expression is dependent on the health, general condition or vigor of the signaler are inherently honest expressions of that individual's quality. For a high-quality signaler, a 'bigger' signal is a smaller handicap (less costly, or more affordable) than it is for a low-quality signaler, thus 'big' signals are reliable signals of signaler quality. One of the clearest demonstrations of honesty in an epigamic signal was carried out by Petrie and her colleagues on that poster animal for epigamic signaling, the peacock. Petrie and colleagues demonstrated that in their peacock population, females preferred a mate with more eyespots in his feather train (whether the difference was natural, or produced by experimental manipulation), and that females mated with males with more eyespots had more young surviving to a year of age than females mated to males with fewer eyespots (Petrie et al., 1991; Petrie, 1994; Petrie and Halliday, 1994). Although the generality of these results has been questioned by studies on other populations (Takahashi et al., 2008; Dakin and Montgomerie, 2011), the example provides a clear illustration of the predictions generated by the handicap principle, and how they should be tested.

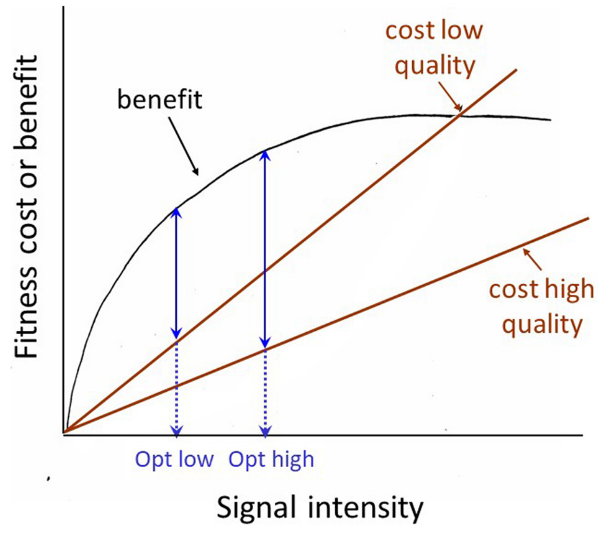


FIGURE 2 | Johnstone's graphical model of the Handicap principle. The basic assumption is that it costs a high-quality signaler less to signal at its optimum level than it costs a low-quality signaler to signal at that level. The optimum or equilibrium level (where the difference between the costs and benefits of signaling are greatest) for the low quality signaler is lower (opt low) than that for the high-quality signaler (opt high). Thus the signaling level is a reliable indicator of signaler quality.

The handicap principle should maintain some degree of honesty in any signaling system where signaler and receiver have non-identical interests, such as virtually all mating and agonistic contexts. A low-quality individual can only 'lie' by diverting energy into signal development and expression that it needs for maintenance, and so as Searcy and Nowicki (2005) succinctly put it, lying becomes more costly than signaling honestly. Searcy and Nowicki suggest that 'reliable' is a better word here than 'honest,' for several reasons. First, as with reliability testing in science and elsewhere, we understand that although perfect reliability is unattainable, partial reliability may be good enough. In contrast, 'honesty' is generally taken to mean absolute honesty. Second, reliability of a signal is empirically measurable. Thus instead of debating whether an animal signal is informative or not, we can measure if it predicts something important about the present state of affairs or future events. Thus for example, in an agonistic situation a 'threat signal' should predict subsequent escalation, and the strongest 'threat' signal should predict attack (Searcy and Beecher, 2009).

Summing Up: Two Perspectives

Historically, the Information Transmission and Manipulation views of animal communication systems have been presented as in opposition. I suggest that in fact they are simply different perspectives on the same process. Once we give the receiver agency, and accept that manipulation is a two-way or reciprocal process in animal communication, we see that the two views have more in common than was at first thought. This rapprochement is nicely captured in the evolution of Dawkins and Krebs's papers on the topic. In their original paper, Dawkins and Krebs (1978) focused on signalers and argued that "natural selection favors [signalers] who successfully manipulate [receivers] whether or not this is to the advantage of the manipulated individuals." However, 6 years later in a follow-up paper (Krebs and Dawkins, 1984) they expanded their view to include receiver interests, noting that receivers would be favored to resist manipulation and to attempt to "read the minds" of signalers. Finally, Krebs (1991), discussing Zahavi's handicap principle, concluded that the manipulation and honest signaling views are probably not incompatible: "Dawkins and Krebs (1978) discussed a coevolutionary process without specifying an end point, whereas Zahavi was concerned mainly with the end-point itself, so it is possible to imagine an evolutionary arms race of manipulation and sales resistance which end up with honest signaling" (Krebs, 1991, p. 67).

Figure 3 is a schematic representation of what I will call the Reciprocal Manipulation view. It shows communication taking place on a battleground in which signaler and receiver are each selected to manipulate the other, the battle being settled in the long run with the compromise of mostly-honest (reliable) signals. The "management-assessment" theory of Owings and Morton (1997, 1998) is quite similar to the Reciprocal Manipulation view. Their theory captures the dynamics of signalers attempting to manage receivers and receivers assessing signalers. In their words "the process of assessment is more active than has been generally recognized, and is responsible for the 'informational' couplings between individuals" (1997, p. 359). However, receivers do more than just assess signalers, they manipulate them as well, requiring them to signal in the first place, and requiring a relatively honest signal as a prerequisite for responding to the signal. The Reliable Signaling view of Searcy and Nowicki (2005) is essentially identical to the Reciprocal Manipulation view, with the superficial difference that the former focuses on the information transmission aspect (reliable signaling) while the latter focuses on the manipulation aspect (the conflicting motivations of signaler and receiver).

The Reciprocal Manipulation and Information Transmission views each seem most helpful in different circumstances (**Table 3**). Where the interests and thus motivations of the two parties differ, the Reciprocal Manipulation highlights the clash. In contrast, where the interests and motivations of the two parties are more in line, the Information Transmission viewpoint focuses on the essence of the interaction. Indeed, where the overlap of sender and receiver interests is considerable, as for example between related individuals, or mates caring for offspring, or individuals in a social group where individuals are strongly interdependent, reliable, mutually beneficial signals will be favored. But even where the interests of sender and receiver are partially opposed, selection acting on both parties will move them to the region where both parties benefit on average, and signals will still be reliable, if less so. This game theory dynamic has been clearly laid out elsewhere (Maynard Smith and Harper, 2003; Godfrey-Smith, 2013).

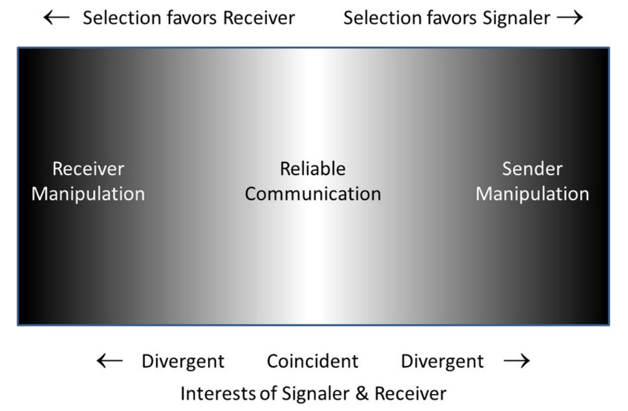


FIGURE 3 | Schematic suggesting the opposing pressures favoring signaler over receiver or vice-versa. Where interests of signaler and receiver are coincident or nearly so (light gray to white) reliable communication will occur. At the extremes of the space (darker), where interests of one or the other of the two parties predominates, signaling will be disfavored. In the intermediate (gray) region, one party may benefit more than the other, but signaling may still be 'reliable enough.'

I believe that the clash between these views of animal communication has ultimately led us to a clearer view of animal communication systems than the original human-oriented information transmission view. Most animal communication systems are somewhere on the continuum from pure manipulation to pure communication, from arms race (where sender and receiver have different interests, each selected to behave so as to benefit themselves) to pure information transmission (where sender and receiver have identical interests, and where signals benefit or cost both parties in the same way or to the same degree). A fuller development of these ideas can be found in Beecher (2020).

In conclusion, I have argued that we should expect that natural communication systems will generally be reliable, even if not perfectly honest, with signaler and receiver both benefiting on average. However, returning to the main theme of this paper, there is no reason to expect such systems to blossom into simple languages unless signalers and receivers have identical or near-identical interests, and if the ecological selective context requires strong cooperation. There are cognitive prerequisites as well – otherwise one might predict that honeybees should have a simple language – but the brake on the evolution to language-like signaling systems in species with the requisite cognitive capacity is provided by the generally divergent interests of signaler and receiver. Otherwise, bonobos, dolphins and some other vertebrates who seem to have the necessary cognitive prerequisites would have a more language-like natural communication systems than they do.

WHY ARE THERE NO NATURAL LANGUAGE SYSTEMS IN ANIMALS?

Research on teaching animals simple human language indicate that at least some animals appear to have the cognitive capacity to decode language or language-like expressions. Herman's dolphins could comprehend a sign language command such as "take the ball to the hoop" and to distinguish it from a similar but syntactically different command like "take the hoop to the ball" (Herman, 2010). Kanzi the bonobo could respond correctly to novel verbal commands such as "Can you put the pine needles in the refrigerator?" (Savage-Rumbaugh et al., 1993). Pepperberg (1981, 1987) and Pailian et al. (2020) have shown that African gray parrots can follow verbal directions to solve difficult problems, including some that challenge humans. Yet despite having the apparent capacities, at least to some extent, no non-human animal uses even a rudimentary language in its day-to-day existence. This includes groups like the songbirds that seem to have a crucial design feature, the learning and cultural transmission of a complex set of vocal signals. Some animals

TABLE 3 | Differences between reciprocal manipulation and information transmission perspectives.

	Perspective	
	Reciprocal manipulation	**Information transmission**
Focus on which aspect of the coevolutionary process?	On the process itself	On the end point of the process
Most useful when sender and receiver interests are:	Divergent	Coincident
Focus on what variable?	Differing motivations of sender and receiver	Information transmitted from sender to receiver

appear to be smart enough, or capable enough to handle a simple language, but we have yet to discover an animal communication system – in nature – that rises to this level. Thus it appears that some missing element other than cognitive or motor limitations has blocked language evolution in non-human animals. Although it is possible that yet some other cognitive limitation has not been clearly identified (Hauser et al., 2002; Pinker and Jackendoff, 2005), I focus in this final section on a candidate for the missing element that is not purely a cognitive mechanism.

A clue as to the missing element comes from the honeybee 'dance language.' Despite a relatively simple nervous system, honeybees are able not only to transmit precise information about events in the external world, but also to use this system in two very different contexts (when talking about the location of desirable food sources or about the location of suitable hive sites). The key ingredient for the evolution of this system, I would argue, is zero conflict of interest between sender and receiver. Both scout and recruit are sister sterile workers and they are both working to feed sisters and brothers slated to be future reproductives. Humans also evolved in a social system featuring extraordinary levels of cooperation, but significantly this cooperation was not restricted to close relatives, as it is in the honeybees and other social insects, ruling out kin selection as a sufficient explanation (but see Fitch, 2004).

I will reframe the question from "why not them?" to the question of "why us" (phrasing suggested by Hrdy, 2009)? How did the human animal become the one species to evolve language? As I argued in the previous section, the field has arrived at a consensus concerning the factors that shape animal communication systems: the pressure for sender and receiver each to shape the interaction to its benefit inevitably both stimulates and constrains the evolution of the communication system. Very unusual circumstances are required for a true language system to evolve. Three essential conditions have to be met. First, the species must have the underlying cognitive capacity. Honeybees may lack this, but some other animals may have it. Second, and this is the clue provided by honeybees, sender and receiver must have identical or near identical interests. Third, individuals must have a compelling need to transmit information across multiple contexts. These are precisely the conditions that existed in pre-human and early human hunter-gatherer societies, the context in which humans and our hominid precursors spent some 95% of our evolutionary history. The description of the prototypical hunter-gatherer society that follows is based on information from a number of sources (including Boehm, 1999; Bowles, 2006; Hrdy, 2009; Hill K. et al., 2011; Knight and Power, 2011; Lee, 2018).

Our hunter-gather ancestors lived in small social groups where individuals were strongly interdependent, and cooperation across multiple contexts was essential for survival. Most highly cooperative animal societies such as the eusocial insects are typically just very large families, but the human hunter-gatherer societies we know – and which we assume to be typical of the ancestral type – consisted of members of several kin lines. Thus human societies then – and now as well – required extensive cooperation among unrelated individuals. Humans are the supreme cooperators in the animal world, but because this cooperation is not supported by high kin relatedness, it has to withstand a strong undercurrent of individual competition. We sometimes lose sight of the human affinity for within-group cooperation because of its paradoxical coexistence with intense between-group competition and tribalism. Irreconcilable conflicts within ancestral hunter-gatherer groups surely occurred, but were often resolved by individuals leaving one group for another (hunter-gatherer societies being classic examples of fission-fusion societies).

Students of human evolution, while differing as to what were the key selective contexts, or the key adaptations, all agree that human evolution has been characterized by remarkable levels of within-group cooperation among unrelated individuals, on a scale not seen in any non-human animal. Several contexts stand out as crucial for the high level of cooperation found in hunter-gather societies. They begin, of course, with hunting and gathering. Effective group hunting (usually done by men) requires sharing of information about distant prey and discussion of strategies for capturing prey. In essentially the same way, gathering of plants and fruits (usually done by women) requires the ability to track the growing schedules and locations of many plants and fruits in the area and the ability to discuss and coordinate foraging activities efficiently. Furthermore, hunter-gatherer societies periodically have to pick up and move to a new, more abundant locale. These moves require discussion and group consensus, with input from all parties, especially older, more experienced men and women.

A second, equally important axis of cooperation is child-raising. Humans are unique among primates in the time and cost required to raise an offspring. Humans solved this problem by involving the whole group in the process. Hrdy (2009) has pointed out that this pattern of cooperative breeding sets humans apart from the exclusive mother-centered parenting of our closest relatives, the great apes. In these early human societies, many individuals played a role in the cooperative care. For starters, the whole group participated in that food brought back to the camp was typically shared among all individuals, without reference to their role in procuring the food. Then unlike most mammals, the father participated in child care alongside the mother. Other relatives were routinely involved in direct child care, especially older siblings and grandparents, often aunts and uncles too, and sometimes non-relatives as well.

Finally, within-group cooperation is essential for success in between-group competition, warfare in particular. This aspect of our hunter-gather heritage is strongly debated in anthropology. Using the terms of Lee (2018), the Peaceful school views significant inter-group competition as not beginning until the Agricultural era, when property gave humans something to fight over. The Bellicose school (e.g., Kelly, 2000; Gat, 2015) believes inter-group competition dates further back in our evolutionary past. But whenever it started, warfare would certainly promote adaptations for within-group cooperation.

In recent years various investigators have proposed key adaptations that may have allowed human societies to achieve this high level of cooperation in the absence of the glue of a very high level of kinship. Although there is not complete agreement as to which of these adaptations were most crucial, taken together they coalesce into a suite of psychological adaptations that promote prosocial within-group interactions within a context of near-complete interdependence. Indeed, Tomasello et al. (2012) have dubbed this the Interdependence hypothesis. The specific adaptations include: shared intentionality (Tomasello et al., 2005), egalitarianism (Boehm, 1999), social learning and communication (Herrmann et al., 2007), intersubjectivity and empathy (Hrdy, 2009), moral intuitions (Haidt, 2012), adaptations for teaching and receiving teaching, and thus cultural transmission (Sterelny, 2012; Henrich, 2016; Whiten, 2017), proactive aggression (Wrangham, 2018) and self-domestication (Wrangham, 2019). These adaptations of our social mind appear to be what set us apart from the other great apes, who it has been argued are otherwise just as cognitively advanced (Herrmann et al., 2007). This suite of adaptations has enabled us to live in complex, cooperative societies. Despite our equally extraordinary proactive (deliberate and planned) aggressive tendencies, directed typically at out-groups, as in wars, pogroms, crusades and the like (Wrangham, 2018), no other social animal has achieved the level of within-group docility and cooperation without high within-group relatedness that is found in the human species. I note that Knight (2018) has an advanced an argument similar to the one I have presented here.

Language unquestionably represents the pinnacle of evolved animal communication systems, and as noted at the beginning of this section, attempts to teach language to animals have not significantly changed this view. Language is often given pride of place in human evolution. In this view the other adaptations mentioned above came only after some form of language was in place. I favor the view of Hrdy (2009), that this may well reverse cause and effect. The evolution of language may have only become possible when the posited unique suite of prosocial, communicative and mind-reading adaptations were in place. The crucial importance of communication in the strongly interdependent social system of early humans would have created this prosocial suite of adaptations, and would have laid the groundwork for evolving a true language.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

Many thanks to editor IP, three reviewers, John Byers, Doug Mock, Trish Schwagmeyer, and Bill Searcy for their very thoughtful reviews of the manuscript.

REFERENCES

Akçay, Ç, Anderson, R. C., Nowicki, S., Beecher, M. D., and Searcy, W. A. (2015a). Quiet threats: soft song as an aggressive signal in birds. *Anim. Behav.* 105, 267–274. doi: 10.1016/j.anbehav.2015.03.009

Akçay, Ç, Campbell, S. E., and Beecher, M. D. (2015b). The fitness consequences of honesty: under-signalers have a survival advantage in song sparrows. *Evolution* 69, 3186–3193. doi: 10.1111/evo.12818

Akçay, Ç, Reed, V. A., Campbell, S. E., Templeton, C. N., and Beecher, M. D. (2010). Indirect reciprocity: song sparrows distrust aggressive neighbors based on eavesdropping. *Anim. Behav.* 80, 1041–1047. doi: 10.1016/j.anbehav.2010.09.009

Akçay, Ç, Tom, M. E., Campbell, S. E., and Beecher, M. D. (2013). Song type matching is an honest early threat signal in a hierarchical animal communication system. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20122517. doi: 10.1098/rspb.2012.2517

Akçay, Ç, Tom, M. E., Holmes, D., Campbell, S. E., and Beecher, M. D. (2011). Sing softly and carry a big stick: signals of aggressive intent in song sparrows. *Anim. Behav.* 82, 377–382. doi: 10.1016/j.anbehav.2011.05.016

Akçay, Ç, Wood, W. E., Searcy, W. A., Templeton, C. N., Campbell, S. E., and Beecher, M. D. (2009). Good neighbour, bad neighbour: song sparrows retaliate against aggressive rivals. *Anim. Behav.* 78, 97–102. doi: 10.1016/j.anbehav.2009.03.023

Anderson, R. C., Searcy, W. A., Peters, S., Hughes, M., DuBois, A. L., and Nowicki, S. (2017). Song learning and cognitive ability are not consistently related in a songbird. *Anim. Cogn.* 20, 309–320. doi: 10.1007/s10071-016-1053-7

Andersson, M. (1982). Female choice selects for extreme tail length in a widowbird. *Nature* 299, 818–820. doi: 10.1038/299818a0

Baker, M. C., and Boylan, J. T. (1995). A catalog of song syllables of indigo and lazuli buntings. *Condor* 97, 1028–1040. doi: 10.2307/1369541

Barker, A. J., Veviurko, G., Bennett, N. C., Hart, D. W., Mograby, L., and Lewin, G. R. (2021). Cultural transmission of vocal dialect in the naked mole-rat. *Science* 371, 503–507. doi: 10.1126/science.abc6588

Barske, J., Schlinger, B. A., Wikelski, M., and Fusani, L. (2011). Female choice for male motor skills. *Proc. R. Soc. Lond., B Biol. Sci.* 278, 3523–3528. doi: 10.1098/rspb.2011.0382

Beecher, M. D. (2017). Birdsong learning as a social process. *Anim. Behav.* 124, 233–246. doi: 10.1016/j.anbehav.2016.09.001

Beecher, M. D. (2020). "Animal communication," in *Oxford Encyclopedia of the History of Psychology*, ed. W. Pickren (Oxford, UK: Oxford University Press).

Beecher, M. D., and Akçay, Ç (2014). "Friends and enemies: how social dynamics shape communication and song learning in song sparrows," in *Animal Behavior*, ed. K. Yakusawa (Santa Barbara, CA: Praeger), 33–61.

Beecher, M. D., Akçay, Ç, and Campbell, S. E. (2020). Birdsong learning is mutually beneficial for tutee and tutor in song sparrows. *Anim. Behav.* 166, 281–288. doi: 10.1016/j.anbehav.2020.05.015

Beecher, M. D., and Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends Ecol. Evol.* 20, 143–149. doi: 10.1016/j.tree.2005.01.004

Beecher, M. D., and Campbell, S. E. (2005). The role of unshared songs in singing interactions between neighbouring song sparrows. *Anim. Behav.* 70, 1297–1304. doi: 10.1016/j.anbehav.2005.03.008

Beecher, M. D., Campbell, S. E., and Burt, J. M. (1994). Song perception in the song sparrow: birds classify by song type but not by singer. *Anim. Behav.* 47, 1343–1351. doi: 10.1006/anbe.1994.1182

Beecher, M. D., Campbell, S. E., Burt, J. M., Hill, C. E., and Nordby, J. C. (2000). Song type matching between neighboring song sparrows. *Anim. Behav.* 59, 21–27. doi: 10.1006/anbe.1999.1276

Beecher, M. D., Nordby, J. C., Campbell, S. E., Burt, J. M., Hill, C. E., and O'Loghlen, A. L. (1997). "What is the function of song learning in songbirds?," in *Communication. Perspectives in Ethology*, Vol. 12, eds D. H. Owings, M. D. Beecher, and N. S. Thompson (New York, NY: Plenum Press), 77–97. doi: 10.1007/978-1-4899-1745-4_4

Beecher, M. D., Stoddard, P. K., Campbell, S. E., and Horning, C. L. (1996). Repertoire matching between neighbouring song sparrows. *Anim. Behav.* 51, 917–923. doi: 10.1006/anbe.1996.0095

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Cambridge, MA: Harvard University Press.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science* 314, 1569–1572. doi: 10.1126/science.1134829

Bradbury, J. W., and Vehrencamp, S. L. (1998). *The Principles of Animal Communication*. Sunderland, MA: Sinauer Associates.

Burt, J. M., Bard, S. C., Campbell, S. E., and Beecher, M. D. (2002). Alternative forms of song matching in song sparrows. *Anim. Behav.* 63, 1143–1151. doi: 10.1006/anbe.2002.3011

Burt, J. M., Campbell, S. E., and Beecher, M. D. (2001). Song type matching as threat: a test using interactive playback. *Anim. Behav.* 62, 1163–1170. doi: 10.1006/anbe.2001.1847

Byers, B. E. (1996). Messages encoded in the songs of chestnut-sided warblers. *Anim. Behav.* 52, 691–705. doi: 10.1006/anbe.1996.0214

Byers, J., Hebets, E., and Podos, J. (2010). Female mate choice based upon male motor performance. *Anim. Behav.* 79, 771–778. doi: 10.1016/j.anbehav.2010.01.009

Cade, W. (1975). Acoustically orienting parasitoids: fly phonotaxis to cricket song. *Science* 190, 1312–1313. doi: 10.1126/science.190.4221.1312

Carazo, P., and Font, E. (2010). Putting information back into biological communication. *J. Evol. Biol.* 23, 661–669. doi: 10.1111/j.1420-9101.2010.01944.x

Caro, T. M., and Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Q. Rev. Biol.* 67, 151–174.

Carouso-Peck, S., and Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. *Curr. Biol.* 29, 631–636.e633. doi: 10.1016/j.cub.2018.12.026

Catchpole, C. K. (1987). Bird song, sexual selection and female choice. *Trends Ecol. Evol.* 2, 94–97. doi: 10.1016/0169-5347(87)90165-0

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York, NY: Praeger.

Clutton-Brock, T. H., and Albon, S. D. (1979). The roaring of red deer and the evolution of honest advertisement. *Behaviour* 69, 145–170. doi: 10.1163/156853979X00449

Coleman, S. W., Patricelli, G. L., Coyle, B., Siani, J., and Borgia, G. (2007). Female preferences drive the evolution of mimetic accuracy in male sexual displays. *Biol. Lett.* 3, 463–466. doi: 10.1098/rsbl.2007.0234

Collins, S. A. (2004). "Vocal fighting and flirting: the functions of birdsong," in *Nature's Music: The Science of Birdsong*, ed. P. M. H. Slabbekoorn (New York, NY: Academic Press), 39–79. doi: 10.1016/b978-012473070-0/50005-0

Dakin, R., and Montgomerie, R. (2011). Peahens prefer peacocks displaying more eyespots, but rarely. *Anim. Behav.* 82, 21–28. doi: 10.1016/j.anbehav.2011.03.016

Davies, N. B., and Halliday, T. R. (1978). Deep croaks and fighting assessment in toads Bufo bufo. *Nature* 274, 683–685. doi: 10.1038/274683a0

Dawkins, R., and Krebs, J. R. (1978). "Animal Signals: information or manipulation?," in *Behavioural Ecology: An Evolutionary Approach*, eds J. R. Krebs and N. B. Davies (Oxford: Blackwell), 282–309.

Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567

Elias, D. O., Maddison, W. P., Peckmezian, C., Girard, M. B., and Mason, A. C. (2012). Orchestrating the score: complex multimodal courtship in the *Habronattus coecatus* group of *Habronattus* jumping spiders (Araneae: Salticidae). *Biol. J. Linn. Soc.* 105, 522–547. doi: 10.1111/j.1095-8312.2011.01817.x

Fisher, J. B. (1954). "Evolution and bird sociality," in *Evolution as Process*, eds J. Huxley, A. C. Hardy, and E. B. Ford (London: Allen & Unwin), 71–83.

Fitch, W. T. (2004). "Evolving honest communication systems: kin delection and 'mother tongues'," in *Evolution of Communication Systems: A Comparative Approach*, eds D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 275–296.

Fitch, W. T., and Jarvis, E. (2013). "Birdsong and other animal models for human speech, song, and vocal learning," in *Language, Music, and the Brain: A Mysterious Relationship*, ed. M. A. Arbib (Cambridge, MA: MIT Press).

Garland, E. C., and McGregor, P. K. (2020). Cultural transmission, evolution, and revolution in vocal displays: insights from bird and whale song. *Front. Psychol.* 11:544929. doi: 10.3389/fpsyg.2020.544929

Gat, A. (2015). Proving communal warfare among hunter-gatherers: the quasi-rousseauan error. *Evol. Anthropol.* 24, 111–126. doi: 10.1002/evan.21446

Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207. doi: 10.1038/nature04675

Godfrey-Smith, P. (2013). "Information and influence in sender-receiver models, with applications to animal behavior," in *Animal Communication Theory: Information and Influence*, ed. U. E. Stegmann (Cambridge: Cambridge University Press).

Gould, J. L. (1975). Honey bee recruitment: the dance-language controversy. *Science* 189, 685–693. doi: 10.1126/science.1154023

Grafen, A. (1990). Biological signals as handicaps. *J. Theor. Biol.* 144, 517–546. doi: 10.1016/s0022-5193(05)80088-8

Green, S. (1975). "Variation of vocal pattern with social situation in the Japanese macaque (*Macaca fuscata*): a field study," in *Primate Behavior*, ed. L. A. Rosenblum (New York, NY: Academic Press), 1–102. doi: 10.1016/b978-0-12-534004-5.50006-3

Green, S., and Marler, P. (1979). "The analysis of animal communication," in *Social Behavior and Communication*, eds P. Marler and G. Vandenbergh (New York, NY: Plenum), 73–158. doi: 10.1007/978-1-4615-9116-0_3

Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York, NY: Pantheon.

Hauser, M. D. (2000). A primate dictionary? Decoding the function and meaning of another species' vocalizations. *Cogn. Sci.* 24, 445–475. doi: 10.1207/s15516709cog2403_5

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

Hebets, E. A., and Uetz, G. W. (1999). Female responses to isolated signals from multimodal male courtship displays in the wolf spider genus*Schizocosa*(Araneae: Lycosidae). *Anim. Behav.* 57, 865–872. doi: 10.1006/anbe.1998.1048

Henrich, J. (2016). *The Secret of our Success: How Culture is Driving Human Evolution, Domesticating our Species, and Making us Smarter*. Princeton, NJ: Princeton University Press.

Herman, L. M. (2010). What laboratory research has told us about dolphin cognition. *Int. J. Comp. Psychol.* 23, 310–330.

Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366. doi: 10.1126/science.1146282

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Curr. Anthrop.* 14, 5–24.

Hill, C. E., Akçay, Ç, Campbell, S. E., and Beecher, M. D. (2011). Extrapair paternity, song and genetic quality in song sparrows. *Behav. Ecol.* 22, 73–81. doi: 10.1093/beheco/arq171

Hill, K., Walker, R. S., Božičević, M., Eder, J., Headland, T., Hewlett, B., et al. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science* 331, 1286–1289. doi: 10.1126/science.1199071

Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 88–111.

Horning, C. L., Beecher, M. D., Stoddard, P. K., and Campbell, S. E. (1993). Song perception in the song sparrow: importance of different parts of the song in song type classification. *Ethology* 94, 46–58. doi: 10.1111/j.1439-0310.1993.tb00546.x

Hrdy, S. B. (2009). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge, MA: Harvard University Press.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* 28, 60–65. doi: 10.1016/j.conb.2014.06.010

Johnstone, R. A. (1997). "The evolution of animal signals," in *Behavioural Ecology: An Evolutionary Approach*, eds J. Krebs and R. Davies (Oxford: Blackwell), 155–178.

Keagy, J., Savard, J.-F., and Borgia, G. (2009). Male satin bowerbird problem-solving ability predicts mating success. *Anim. Behav.* 78, 809–817. doi: 10.1016/j.anbehav.2009.07.011

Kelly, R. (2000). *Warless Societies and the Origin of War*. Ann Arbor: Uinversity of Michigan Press.

Knight, C. (2018). "Pressure for trust-based efficiency shaped the evolution of language," in *The Evolution of Language: Proceedings of the 12th International Conference (EVOLANGXII)*, eds C. Christine, F. Molly, L. Hannah, M. Luke, R. Andrea, and V. Tessa.

Knight, C., and Power, C. (2011). "Social conditions for the evolutionary emergence of language," in *Handbook of Language Evolution*, ed. K. G. M. Tallerman (Oxford: Oxford University Press), 346–349.

Knornschild, M. (2014). Vocal production learning in bats. *Curr. Opin. Neurobiol.* 28, 80–85. doi: 10.1016/j.conb.2014.06.014

Krebs, J. R. (1991). "Animal communication: ideas derived from Tinbergen's activities," in *The Tinbergen Legacy*, eds M. D. Dawkins, T. R. Halliday, and R. Dawkins (London: Chapman & Hall), 60–74. doi: 10.1007/978-0-585-35156-8_5

Krebs, J. R., and Davies, R. (1978). *Behavioural Ecology: An Evolutionary Approach.* Oxford: Blackwell.

Krebs, J. R., and Dawkins, R. (1984). "Animal signals: mind-reading and manipulation," in *Behavioural Ecology: An Evolutionary Approach*, 2nd Edn, eds J. R. Krebs and R. Davies (Oxford: Blackwell), 380–402.

Kroodsma, D., Hamilton, D., Sánchez, J. E., Byers, B. E., Fandiño-Mariño, H., Stemple, D. W., et al. (2013). Behavioral evidence for song learning in the suboscine bellbirds (Procnias spp.; Cotingidae). *Wilson J. Ornithol.* 125, 1–14. doi: 10.1676/12-033.1

Kroodsma, D. E. (1977). Correlates of song organization among North American wrens. *Am. Nat.* 111, 995–1008. doi: 10.1086/283228

Kroodsma, D. E. (1988). "Contrasting styles of song development and their consequences among passerine birds," in *Evolution and Learning*, eds R. C. Bolles and M. D. Beecher (Hillsdale, NJ: Lawrence Erlbaum Associates), 157–184.

Langmore, N. E. (1998). Functions of duet and solo songs of female birds. *Trends Ecol. Evol.* 13, 136–140. doi: 10.1016/s0169-5347(97)01241-x

Lee, R. L. (2018). Hunter-gatherers and human evolution: new light on old debates. *Annu. Rev. Anthrop.* 47, 513–531. doi: 10.1146/annurev-anthro-102116-041448

Levin, R. N. (1996). Song behaviour and reproductive strategies in a duetting wren, *Thryothorus nigricapillus*: II. Playback experiments. *Anim. Behav.* 52, 1107–1117. doi: 10.1006/anbe.1996.0258

Logue, D. M., and Gammon, D. E. (2004). Duet songs and sex roles during territory defence in a tropical bird, the black-bellied wren. *Anim. Behav.* 68, 721–731. doi: 10.1016/j.anbehav.2003.10.026

Lorenz, K. Z. (1966). The psychobiological approach: methods and results–Evolution of ritualization in the biological and cultural spheres. *Philos. Trans. R. Soc. Lond. B Biol. Sci* 251, 273–284. doi: 10.1098/rstb.1966.0011

MacDougall-Shackleton, S. A. (1997). Sexual selection and the evolution of song repertoires. *Curr. Ornithol.* 14, 81–124. doi: 10.1007/978-1-4757-9915-6_3

Marler, P. (1961). The logical analysis of animal communication. *J. Theor. Biol.* 1, 295–317. doi: 10.1016/0022-5193(61)90032-7

Marler, P. (1968). "Visual systems," in *Animal Communication*, ed. T. A. Sebeok (Bloomington, IN: Indiana University Press), 103–126.

Marler, P. (1970). Birdsong and speech development: could there be parallels? *Am. Sci.* 58, 669–673.

Marler, P., and Pickert, R. (1984). Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Anim. Behav.* 32, 673–689. doi: 10.1016/s0003-3472(84)80143-8

Marler, P., and Tamura, M. (1964). Culturally transmitted patterns of vocal behavior in sparrows. *Science* 146, 1483–1486. doi: 10.1126/science.146.3650.1483

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

McDonald, M. V. (1989). Function of song in Scott's seaside sparrow. *Anim. Behav.* 38, 468–485. doi: 10.1016/s0003-3472(89)80040-5

Medvin, M. B., Stoddard, P. K., and Beecher, M. D. (1993). Signals for parent-offspring recognition: a comparative analsysis of the begging calls of cliff swallows and barn swallows. *Anim. Behav.* 45, 841–850. doi: 10.1006/anbe.1993.1105

Nowicki, S., and Searcy, W. A. (2014). The evolution of vocal learning. *Curr. Opin. Neurobiol.* 28, 48–53. doi: 10.1016/j.conb.2014.06.007

Otte, D. (1974). Effects and functions in the evolution of signaling systems. *Annu. Rev. Ecol. Syst.* 5, 385–417. doi: 10.1146/annurev.es.05.110174.002125

Otter, K. A., Mckenna, A., LaZerte, S. E., and Ramsay, S. M. (2020). Continent-wide shifts in song dialects of white-throated sparrows. *Curr. Biol.* 30, 1–5. doi: 10.1016/j.cub.2020.05.084

Owings, D. H., and Morton, E. S. (1997). "The role of information in communication: an assessment/management approach," in *Communication*, eds D. H. Owings, M. D. Beecher, and N. S. Thompson (New York, NY: Plenum Press), 359–390. doi: 10.1007/978-1-4899-1745-4_12

Owings, D. H., and Morton, E. S. (1998). *Animal Vocal Communication: A New Approach*. Cambridge: Cambridge University Press.

Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biol. Philos.* 25, 755–780. doi: 10.1007/s10539-010-9224-4 Corpus ID: 13443399,

Pailian, H., Carey, S. E., Halberda, J., and Pepperberg, I. M. (2020). Age and species comparisons of visual mental manipulation ability as evidence for its development and evolution. *Sci. Rep.* 10:7689. doi: 10.1038/s41598-020-64666-1

Penn, D. J., and Számadó, S. (2018). The handicap principle: how an erroneous hypothesis became a scientific principle. *Biol. Rev.* 95, 267–290. doi: 10.1111/brv.12563

Pepperberg, I. M. (1981). Functional vocalizations by an African grey parrot (*Psittacus erithacus*). *Z. Tierpsychol.* 55, 139–160. doi: 10.1111/j.1439-0310.1981.tb01265.x

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. *Anim. Learn. Behav.* 15, 423–432. doi: 10.3758/bf03205051

Petrie, M. (1994). Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* 371, 598–599. doi: 10.1038/371598a0

Petrie, M., Halliday, T., and Sanders, C. (1991). Peahens prefer peacocks with elaborate trains. *Anim. Behav.* 41, 323–331. doi: 10.1016/s0003-3472(05)80484-1

Petrie, M., and Halliday, T. R. (1994). Experimental and natural changes in the peacock's (*Pavo cristatus*) train can affect mating success. *Behav. Ecol. Sociobiol.* 35, 213–217. doi: 10.1007/bf00167962

Pinker, S., and Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition* 95, 201–236. doi: 10.1016/j.cognition.2004.08.004

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5:13220. doi: 10.1038/srep13220

Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Hiebert, S. M., Smith, J. N. M., Stoddard, P. K., et al. (2004). Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Anim. Behav.* 68, 1055–1063. doi: 10.1016/j.anbehav.2004.07.003

Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Hiebert, S. M., Smith, J. N. M., Stoddard, P. K., et al. (2005). Fitness correlates of song repertoire size in free-living song sparrows *(Melospiza melodia). Am. Nat.* 165, 299–310. doi: 10.2307/3473407

Rendall, D., Owren, M. J., and Ryan, M. J. (2009). What do animal signals mean? *Anim. Behav.* 78, 233–240. doi: 10.1016/j.anbehav.2009.06.007

Rice, J. O., and Thompson, W. L. (1968). Song development in the indigo bunting. *Anim. Behav.* 16, 462–469. doi: 10.1016/0003-3472(68)90041-9

Ryan, M. J., and Rand, A. S. (1990). The sensory basis of sexual selection for complex calls in the Túngara frog, *Physalaemus pustulosus* (sexual selection for sensory exploitation). *Evolution* 44, 305–314. doi: 10.1111/j.1558-5646.1990.tb05200.x

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., and Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monogr. Soc. Res. Child Dev.* 58, 1–222.

Schürch, R., and Ratnieks, F. W. (2015). The spatial information content of the honey bee waggle dance. *Front. Ecol. Evol.* 3:22. doi: 10.3389/fevo.2015.00022

Scott-Phillips, T. C. (2008). Defining biological communication. *J. Evol. Biol.* 21, 387–395. doi: 10.1111/j.1420-9101.2007.01497.x

Searcy, W. A. (1984). Song repertoire size and female preferences in song sparrows. *Behav. Ecol. Sociobiol.* 14, 281–286. doi: 10.1007/bf00299499

Searcy, W. A., Akçay, Ç, Nowicki, S., and Beecher, M. D. (2014). Aggressive signaling in song sparrows and other songbirds. *Adv. Study Behav.* 46, 89–125. doi: 10.1016/b978-0-12-800286-5.00003-1

Searcy, W. A., and Beecher, M. D. (2009). Song as an aggressive signal in songbirds. *Anim. Behav.* 78, 1281–1292. doi: 10.1016/j.anbehav.2009.08.011

Searcy, W. A., and Nowicki, S. (2005). *The Evolution of Animal Communication*. Princeton, NJ: Princeton University Press.

Searcy, W. A., and Yasukawa, K. (1996). "Song and female choice," in *Ecology and Evolution of Acoustic Communication in Birds*, eds D. E. M. Kroodsma and H. Edward (London: Cornell University Press), 454–473.

Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Zuberbühler, K., and Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Anim. Behav.* 80, 3–8. doi: 10.1016/j.anbehav.2010.04.012

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1990). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Smith, W. J. (1997). "The behavior of communicating, after twenty years," in *Communication*, eds D. H. Owings, M. D. Beecher, and N. S. Thompson (New York, NY: Plenum), 7–53. doi: 10.1007/978-1-4899-1745-4_2

Stein, A. C., and Uy, J. A. C. (2006). Plumage brightness predicts male mating success in the lekking golden-collared manakin, *Manacus vitellinus*. *Behav. Ecol.* 17, 41–47. doi: 10.1093/beheco/ari095

Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, MA: The MIT Press.

Stoddard, P. K., and Beecher, M. D. (1983). Parental recognition of offspring in the cliff swallow. *Auk* 100, 795–799. doi: 10.1093/auk/100.4.795

Stoeger, A., and Manger, P. (2014). Vocal learning in elephants: neural bases and adaptive context. *Curr. Opin. Neurobiol.* 28, 101–107. doi: 10.1016/j.conb.2014.07.001

Struhsaker, T. T. (1967). "auditory communication among vervet monkeys *(Cercopithecus aethiops)*," in *Social Communication Among Primates*, ed. S. A. Altmann (Chicago, IL: Chicago University Press), 281–324.

Studdert-Kennedy, M. (1998). "The particulate origins of language generativity: from syllable to gesture," in *Approaches to the Evolution of Language: Social and Cognitive Bases*, eds J. Hurford, M. Studdert-Kennedy, and C. Knight (Cambridge: Cambridge University Press), 202–221.

Takahashi, M., Arita, H., Hiraiwa-Hasegawa, M., and Hasegawa, T. (2008). Peahens do not prefer peacocks with more elaborate trains. *Anim. Behav.* 75, 1209–1219. doi: 10.1016/j.anbehav.2007.10.004

Templeton, C. N., Akçay, Ç, Campbell, S. E., and Beecher, M. D. (2012). Soft song is a reliable signal of aggressive intent in song sparrows. *Behav. Ecol. Sociobiol.* 66, 1503–1509. doi: 10.1007/s00265-012-1405-5

Thompson, W. L. (1969). Song recognition by territorial male buntings (Passerin A). *Anim. Behav.* 17, 658–663. doi: 10.1016/S0003-3472(69)80008-4

Thompson, W. L. (1970). Song variation in a population of indigo buntings. *Auk* 87, 58–71. doi: 10.2307/4083658

Tinbergen, N. (1952). Derived activities: their causation, biological significance, origin and emancipation during evolution. *Q. Rev. Biol.* 27, 1–32. doi: 10.1086/398642

Tinbergen, N. (1964). "The evolution of signaling devices," in *Social Behavior and Evolution Among Vertebrates*, ed. W. Etkin (Chicago, IL: University of Chicago Press), 206–230.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/s0140525x05000129

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., and Herrmann, E. (2012). Two key steps in the evolution of human cooperation: the interdependence hypothesis. *Curr. Anthropol.* 53, 673–692. doi: 10.1086/668207

van Heijningen, C. A. A., de Visser, J., Zuidema, W., and ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20538. doi: 10.1073/pnas.0908113106

Wacewicz, S., and Żywiczyński, P. (2015). Language evolution: why Hockett's design features are a non-starter. *Biosemiotics* 8, 29–46. doi: 10.1007/s12304-014-9203-2

Walcott, C., Mager, J. N., and Piper, W. (2006). Changing territories, changing tunes: male loons, *Gavia immer*, change their vocalizations when they change territories. *r Anim. Behav.* 71, 673–683. doi: 10.1016/j.anbehav.2005.07.011

Whiten, A. (2017). Social learning and culture in child and chimpanzee. *Annu. Rev. Psychol.* 68, 129–154. doi: 10.1146/annurev-psych-010416-044108

Wiley, R. H. (2013). "Communication as a transfer of information: measurement, mechanism, and meaning," in *Animal Communication Theory: Information and Influence*, ed. U. Stegmann (Cambridge: Cambridge University Press), 113–129. doi: 10.1017/cbo9781139003551.007

Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Wilson, E. O. (1975). *Sociobiology*. Cambridge, MA: Harvard University Press.

Wrangham, R. W. (2018). Two types of aggression in human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 115, 245–253. doi: 10.1073/pnas.1713611115

Wrangham, R. W. (2019). Hypotheses for the evolution of reduced reactive aggression in the context of human self-domestication. *Front. Psychol.* 10:1914. doi: 10.3389/fpsyg.2019.01914

Wright, T. F., and Dahlin, C. R. (2018). Vocal dialects in parrots: patterns and processes of cultural evolution. *Emu* 118, 50–66. doi: 10.1080/01584197.2017.1379356

Zahavi, A. (1975). Mate selection- a selection for handicap. *J. Theor. Biol.* 53, 205–214. doi: 10.1016/0022-5193(75)90111-3

Zuidema, W., and de Boer, B. (2009). The evolution of combinatorial phonology. *J. Phon.* 37, 125–144. doi: 10.1016/j.wocn.2008.10.003

Zuk, M., Rotenberry, J. T., and Tinghitella, R. M. (2006). Silent night: adaptive disappearance of a sexual signal in a parasitized population of field crickets. *Biol. Lett.* 2, 521–524. doi: 10.1098/rsbl.2006.0539

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

GENERAL COMMENTARY
published: 04 October 2021
doi: 10.3389/fpsyg.2021.722685

Commentary: Why Are No Animal Communication Systems Simple Languages?

Dustin J. Penn[1]* *and Szabolcs Számadó*[2,3]

[1] Department of Interdisciplinary Life Sciences, Konrad Lorenz Institute of Ethology, University of Veterinary Medicine, Vienna, Austria, [2] Department of Sociology and Communication, Budapest University of Technology and Economics, Budapest, Hungary, [3] Center for Social Sciences, Eötvös Loránd Research Network (ELKH), Budapest, Hungary

Keywords: language, human evolution, animal communication, handicap principle, gene-culture co-evolution

OPEN ACCESS

Edited by:
Wei Chen,
Shaoxing University, China

Reviewed by:
Zhanna Reznikova,
Institute of Systematics and Ecology of Animals (RAS), Russia
Gisela - Kaplan,
University of New England, Australia
Michael Charles Corballis,
The University of Auckland, New Zealand
Gillian Forrester,
Birkbeck, University of London, United Kingdom
Slawomir Wacewicz,
Nicolaus Copernicus University in Toruń, Poland

****Correspondence:***
Dustin J. Penn
dustin.penn@vetmeduni.ac.at

Specialty section:
This article was submitted to Evolutionary Psychology, a section of the journal Frontiers in Psychology

Received: *09 June 2021*
Accepted: *07 September 2021*
Published: *04 October 2021*

Citation:
Penn DJ and Számadó S (2021) Commentary: Why Are No Animal Communication Systems Simple Languages? Front. Psychol. 12:722685. doi: 10.3389/fpsyg.2021.722685

A Commentary on

Why Are No Animal Communication Systems Simple Languages?
by Beecher, M. D. (2021). Front. Psychol. 12:602635. doi: 10.3389/fpsyg.2021.602635

> "Language, too, has apparently evolved only in us: that is to say 40 times less often than eyes. It is surprisingly hard to think of 'good ideas' that have evolved only once."–Dawkins (2004), p. 592

Ever since Darwin, scientists have sought to understand the origins of human language, which has been called "the hardest problem in science" (see Christiansen and Kirby, 2003; Számadó and Szathmáry, 2006 for review). It is difficult to understand why humans are the only species that evolved language. Beecher (2021), a pioneer in the study of birdsong, recently considered this problem and the implications of animal communication research for the evolution of language. He did a splendid job describing the "design features" of language vs. other animal communication systems. However, his summary of honest signaling theory is inaccurate, and he overlooked gene-culture co-evolution for explaining language.

Beecher first clarifies the design features that make human language different from the communication systems of other species (i.e., semanticity, nearly infinite information capacity, arbitrariness, evolvability via cultural transmission). He then provides a clear explanation for why the communication of other species lacks the key features of a full-fledged language, and examines birdsong as an example. Song birds have vocal learning, a complex vocal repertoire, and hierarchical structuring of vocal signals, but they do not seem to use these capacities to form different songs to represent different meanings (birds do not transmit different messages with different songs; they do not seem to create new meanings by recombining words into sentences). Birdsong has complexity, but not infinite information capacity. He convincingly argues that animal vocalizations do not function like language.

He then examines the debates over the "fundamental nature of animal communication," and whether animal signals are honest or manipulative; but this summary is not completely accurate. Dawkins and Krebs (1978) challenged the widespread assumption that animal signals are always honest and function to provide information *per se*, and pointed out that their function is to influence conspecifics, which can include persuasion, deception, and manipulation. Some have mistakenly pitted information against influence (Rendall et al., 2009), which are not alternatives; signals can be influential because they inform—or misinform. The main issue at stake here is whether we should expect signals to be honest, but contrary to what Beecher assumes, there is no theoretical justification for the idea that they should be honest *on average*. Animals should evolve sales resistance to avoid manipulation, as he points out, but there are many examples of deception, and the amount deception can be high in theoretical models.

Beecher argues that the handicap principle will maintain some degree of honesty in any signaling system, even though it is "still being subjected to further modification and clarification" (p. 10), and cites our recent review (i.e., Penn and Számadó, 2020). However, we argue that the handicap principle needs to be fully rejected, not modified. The handicap principle is the idea that signals are honest because they are costly to produce, and it assumes that signals are costly is to demonstrate that they are honest. Its logic is circular and non-Darwinian (costly traits evolve despite and not because of their costs), and since theoretical models have shown that signaling costs are not not necessary to maintain honesty (see below), this idea can and should be rejected.

Beecher states that the handicap principle was formalized by Grafen's (1990) model, but this is not a handicap model and it never was. Grafen never showed that signals will be wasteful at the equilibrium, and his equations do not support his handicap interpretation, as they show the necessity of differential marginal costs, and not the necessity of equilibrium costs for honest signaling. Yet, it is not the equilibrium cost of signals (a.k.a. handicaps) that maintain honesty, but the potential cost of cheating (i.e., ratio of marginal benefit to marginal cost for potential cheaters, see Hurd, 1995; Lachmann et al., 2001; Bergstrom et al., 2002). Neither the handicap principle nor Grafen's "main handicap results" are supported by theoretical models or empirical results (see Számadó, 2011; Penn and Számadó, 2020). Beecher suggests that some studies on peacocks provide "a clear illustration of the predictions generated by the handicap principle, and how they should be tested." He is actually referring to Grafen's model rather than the handicap principle, but neither were tested in these studies. Also, Grafen's model is not as general as Beecher assumes, and it is unclear that it can explain signal reliability (e.g., see Nöldeke and Samuelson, 2003).

The handicap principle has generated enormous confusion due to its dressing up a non-Darwinian idea (handicaps) as a Darwinian theory and scientific principle. It misled a generation of biologists into attempting to measure the equilibrium cost of signals, which is uninformative, instead of measuring the marginal costs and marginal *fitness* benefits of signals. Animal communication can be described and analyzed in terms of evolutionary life-history trade-offs, without imposing the confusing language of the handicap paradigm.

Finally, Beecher argues that the key ingredient for the evolution of language is zero conflict of interest between sender and receiver. We humans are exceptional for our capacity to cooperate with strangers (Maynard Smith and Szathmáry, 1995); but we are far from zero conflict. Our species' success seems to be due to our ability to cooperate in large numbers toward common goals, which undoubtedly requires language. It is unclear, however, whether cooperation drove the evolution of language or vice versa. It is difficult to see how one might determine which came first (but see Számadó, 2010), and they likely co-evolved. Darwin (1871) proposed that languages evolve like living organisms, and gene-culture co-evolutionary theory provides important insights into how language and cooperation can influence each other's evolution (e.g., Pinker and Bloom, 1990; Richerson and Boyd, 1999; Számadó and Szathmáry, 2012; see concise summary in Richerson et al., 2021). As language and cooperation evolved, they may have generated positive feedback with each other (e.g., Számadó, 2010). Language likely allowed our ancestors to cooperate, and helped to resolve conflicts by exchanging information, though this includes invented fictions, social constructions, and other imagined realities (Harari, 2014). Honest communication is corruptible, as long as there are conflicts of interest (Dawkins and Guilford, 1991). Yet, for language to function as it does, it need not be completely reliable (and clearly it is not) and conflicts need not be zero; it only needs to facilitate communication.

AUTHOR CONTRIBUTIONS

DP and SS wrote the commentary. All authors contributed to the article and approved the submitted version.

FUNDING

DP acknowledges support from the Austrian Science Fund (AT): P28141-B25 of the Austrian Science Foundation. SS was supported by OTKA grant K132250, and by the European Structural Investment Funds (ESIF) in Hungary (GINOP) 2.3.2-15-2016-00057. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank S. Zala and the five reviewers for comments and R. Hengsberger for assistance.

REFERENCES

Beecher, M. D. (2021). Why are no animal communication systems simple languages? *Front. Psych.* 12:602635. doi: 10.3389/fpsyg.2021.602635

Bergstrom, C. T., Számad,ó, S., and Lachmann, M. (2002). Separating equilibria in continuous signalling games. *Phil. Trans. R. Soc. Lond. B* 357, 1595–1606. doi: 10.1098/rstb.2002.1068

Christiansen, M. H., and Kirby, S. (2003). "Language evolution: the hardest problem in science?," in *Language Evolution,* eds. M. H. Christiansen and S. Kirby (Oxford; New York, NY: Oxford University Press). doi: 10.1093/acprof:oso/9780199244843.001.0001

Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray. doi: 10.5962/bhl.title.2092

Dawkins, M. S., and Guilford, T. (1991). The corruption of honest signalling. *Anim. Behav.* 41, 865–874. doi: 10.1016/S0003-3472(05)80353-7

Dawkins, R. (2004). *The Ancestor's Tale. A Pilgrimage to the Dawn of Evolution*. Boston, MA: Houghton Mifflin.

Dawkins, R., and Krebs, J. R. (1978). "Animal signals: information or manipulation?," in *Behavioural Ecology. An Evolutionary Approach*, eds. J. R. Krebs and N. B. Davies (Oxford: Blackwell), 282–309.

Grafen, A. (1990). Biological signals as handicaps. *J. Theor. Biol.* 144, 517–546. doi: 10.1016/S0022-5193(05)80088-8

Harari, Y. N. (2014). *Sapiens: a Brief History of Humankind*. London: Harvill Secker.

Hurd, P. L. (1995). Communication in discrete action-response games. *J. Theor. Biol.* 174, 217–222. doi: 10.1006/jtbi.1995.0093

Lachmann, M., Számado, S., and Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13189–13194. doi: 10.1073/pnas.231216498

Maynard Smith, J., and Szathmáry, E. (1995). *The Major Transitions in Evolution.* Oxford, New York, Heidelberg: W.H. Freeman - Spektrum.

Nöldeke, G., and Samuelson, L. (2003). Strategic choice handicaps when females seek high male net viability. *J. Theor. Biol.* 221, 53–59. doi: 10.1006/jtbi.2003.3169

Penn, D. J., and Számadó, S. (2020). The Handicap Principle: how an erroneous hypothesis became a scientific principle. *Biol. Rev.* 95, 267–290. doi: 10.1111/brv.12563

Pinker, S., and Bloom, P. (1990). Natural language and natural selection. *Behav. Brain Sci.* 13, 707–727. doi: 10.1017/S0140525X00081061

Rendall, D., Owren, M. J., and Ryan, M. J. (2009). What do animal signals mean? *Anim. Behav.* 78, 233–240. doi: 10.1016/j.anbehav.2009.06.007

Richerson, P. J., and Boyd, R. (1999). Complex societies. the evolutionary origins of a crude superorganism. *Hum. Nat.* 10, 253–289. doi: 10.1007/s12110-999-1004-y

Richerson, P. J., Gavrilets, S., and de Waal, F. B. M. (2021). Modern theories of human evolution foreshadowed by Darwin's Descent of Man. *Science* 372:eaba3776. doi: 10.1126/science.aba3776

Számadó, S. (2010). Pre-hunt communication provides context for the evolution of early human language. *Biol. Theory* 5, 366–382. doi: 10.1162/BIOT_a_00064

Számadó, S. (2011). The cost of honesty and the fallacy of the handicap principle. *Anim. Behav.* 81, 3–10. doi: 10.1016/j.anbehav.2010.08.022

Számadó, S., and Szathmáry, E. (2006). Selective scenarios for the emergence of natural language. *Trends Ecol. Evol.* 21, 555–561. doi: 10.1016/j.tree.2006.06.021

Számadó, S., and Szathmáry, E. (2012). "Evolutionary biological foundations of the origin of language: the co-evolution of language and brain," in *The Oxford Handbook of Language Evolution,* eds. K. R. Gibson and M. Tallerman (Oxford: Oxford University Press), 157–167. doi: 10.1093/oxfordhb/9780199541119.013.0014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

GENERAL COMMENTARY
published: 15 October 2021
doi: 10.3389/fpsyg.2021.763445

Commentary: Why Are no Animal Communication Systems Simple Languages?

*Sławomir Wacewicz**

Center for Language Evolution Studies, Nicolaus Copernicus University in Toruń, Toruń, Poland

Keywords: language evolution, animal communication, conflict of interests, Hockett's design features, platform of trust

A commentary on

Why Are no Animal Communication Systems Simple Languages?
by Beecher, M. D. (2021). Front. Psychol. 12:602635. doi: 10.3389/fpsyg.2021.602635

OPEN ACCESS

Edited by:
Irene M. Pepperberg,
Harvard University, United States

Reviewed by:
Kazuo Okanoya,
The University of Tokyo, Japan
Yosef Prat,
Independent Researcher, Tel Aviv, Israel

****Correspondence:***
Sławomir Wacewicz
wacewicz@umk.pl

Specialty section:
This article was submitted to Evolutionary Psychology, a section of the journal Frontiers in Psychology

Received: *23 August 2021*
Accepted: *20 September 2021*
Published: *15 October 2021*

Citation:
Wacewicz S (2021) Commentary: Why Are no Animal Communication Systems Simple Languages? Front. Psychol. 12:763445. doi: 10.3389/fpsyg.2021.763445

INTRODUCTION

In the title of his paper, Beecher (2021) asks "Why Are No Animal Communication Systems Simple Languages?" In his answer to this question, he identifies two necessary conditions for developing a language-like communication system: "strong cognitive and signal production mechanisms" and a low level of conflict of interests between the communicators. Although this answer is not qualitatively novel, Michael Beecher makes a highly valuable point in stressing this latter condition over the former one: while many animal species have a level of cognitive sophistication that should predispose them to have at least rudimentary languages, such species do not meet the other criterion, that of sufficient alignment of interest. I agree with the essence of this argument, which is still underappreciated in the language evolution literature. However, I am critical of the two main steps of Beecher's proposal, that is the choice of Hockett's design features of language as a starting point, and the presentation of the argument related to the conflict of interests.

COGNITIVE PREREQUISITES FOR LANGUAGE ARE MORE IMPORTANT THAN DESIGN FEATURES

Beecher begins his argument by observing that the communication of birds displays a number of key features adapted from Charles Hockett's (1959) classic set of design features of language. Admittedly, Hockett's system is still the most widely used yardstick of comparing human and non-human communication systems, but after over 60 years it has become theoretically obsolete, and assuming it as a point of departure here is unfortunate for several reasons.

Most importantly, the relevance of the system of design features of language to the main thesis of the paper is only indirect. Beecher's main proposal is that what prevents non-human animals from developing a simple language is a lack of extreme social interdependence, even though many species may have the requisite cognitive abilities. Although I agree with this position, it entails that what truly matters is cognitive abilities rather than design features, which in turn makes Beecher's carefully argued interim conclusion—that many animal communication systems have many of the design features of language—orthogonal to his main argument. This is particularly so that Hockett's system concerns the structural and functional properties of the communicative code and is entirely non-cognitive (which, incidentally, is a strong reason to question its applicability to language evolution research, see in particular Wacewicz and Zywiczyński, 2015). In short, the construction of Beecher's argument calls for addressing cognition directly; instead it is only done via the roundabout route of design features, and the relevant cognitive capacities have not been discussed nor identified. This point is far from trivial, since several cognitive capacities considered as evolutionary preconditions for language have been argued to be uniquely human, such as advanced executive functions (e.g., Adornetti, 2016) or advanced intersubjectivity and triadic bodily mimesis (e.g., Zlatev, 2014).

It should also be noted that Hockett's system misses important features of language that make for the truly crucial differences from the communication of other animals, and while making up for these shortcomings is possible, it often results in terminological problems. As one example, a critically important feature of language is its open-ended semantics (cf. e.g., Arbib, 2012), which depends on the domain-generality of human communication—it is semantically universal in the sense of covering any thematic domain, in contrast to narrowly defined domains for many animals systems, such as food calls or alarm calls. Although Beecher does consider this property, he discusses it under "productivity," which on his account unfortunately conflates three distinct properties of communicative systems: semantic universality, duality of patterning, and productivity in its prototypical meaning of the generative potential of language for structural novelty.

CONFLICT OF INTERESTS AND A PLATFORM OF TRUST

As a second, and central, condition for developing language, Beecher identifies "near-absent" or even "zero" conflict of interest between communicators. Beecher's focus on game-theoretic explanations, with conflict of interest as a key explanatory variable, is certainly valuable and productive; however—as pointed out in another commentary (Penn and Számadó, 2021)—the requirement that communicators have only minimal or zero conflict of interest is both too strong and unrealistic. Contra Beecher, the challenge for explaining language evolution is not how people have got to have near-absent conflict of interests, because they clearly have not: situations involving a different ordering of preferences between human agents are as ubiquitous now as they undoubtedly must have been in our evolutionary past. Rather, the challenge seems to lie in explaining how humans managed to evolve language *in spite of* non-zero conflict of interests, that is, under conditions that signaling theory predicts language-like systems of large-scale, cheap but honest information donation are not evolvable.

A promising direction is to openly admit this dissociation between general behavior and communicative behavior: while humans clearly do not have completely aligned interests, communicatively they behave *as if* they did. A proposal that captures this is a Platform of Trust, which is defined as "a social niche in which large-scale cheap but honest communication is possible because messages tend to be trusted as a default" (Wacewicz and Zywiczyński, 2018, p. 172), but in terms of the explanatory principle of alignment of interest it can be reformulated as "*as-if* alignment of interests between human communicators." Importantly, "Platform of Trust" is neutral on how this communicative alignment of interests arose in human evolution. In other words, it is not an explanatory proposal but an explanatory target, in that it is not itself a scenario of language emergence but rather a necessary constraint for any such scenario. However, having well-defined explanatory targets is conducive to better scenarios, which—crucially—take seriously both the uniqueness of human language and the constraints that signaling theory imposes on *all* systems of communication. This is in line with Beecher's main point, which instead of the already almost universally appreciated cognitive preconditions for language prioritizes looking into the underappreciated factors relevant to the often divergent interests of the communicators. Explanations of the stability of honest communication in human societies in the face of a partial conflict of interests between the communicating humans are likely to refer to mechanisms such as epistemic vigilance (Sperber et al., 2010), gossip (Dunbar, 1996), or reputation formation through indirect reciprocity (Nowak and Sigmund, 2005).

AUTHOR CONTRIBUTIONS

SW wrote the paper.

FUNDING

This research was supported by the Polish National Science Center under grant agreement UMO-2019/34/E/HS2/00248.

REFERENCES

Adornetti, I. (2016). On the phylogenesis of executive functions and their connection with language evolution. *Front. Psychol.* 7:1426. doi: 10.3389/fpsyg.2016.01426

Arbib, M. A. (2012). *How the Brain Got Language: The Mirror System Hypothesis*. Oxford: Oxford University Press. doi: 10.1093/acprof:osobl/9780199896684.001.0001

Beecher, M. D. (2021). Why are no animal communication systems simple languages? *Front. Psychol.* 12:701. doi: 10.3389/fpsyg.2021.02635

Dunbar, R. I. M. (1996). *Grooming, Gossip, and the Evolution of Language.* Cambridge, MA: Harvard University Press.

Hockett, C. F. (1959). Animal 'languages' and human language. *Hum. Biol.* 31, 32–39.

Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437, 1291–1298. doi: 10.1038/nature04131

Penn, D. J., and Számadó, S. (2021). Commentary: why are no animal communication systems simple languages? *Front. Psychol.* 12:722685. doi: 10.3389/fpsyg.2021.722685

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., et al. (2010). Epistemic vigilance. *Mind Lang.* 25, 359–393. doi: 10.1111/j.1468-0017.2010.01394.x

Wacewicz, S., and Zywiczyński, P. (2015). Language evolution: why Hockett's design features are a non-starter. *Biosemiotics* 8, 29–46. doi: 10.1007/s12304-014-9203-2

Wacewicz, S., and Zywiczyński, P. (2018). Language origins: fitness consequences, platform of trust, cooperation, and turn-taking. *Interact. Stud.* 19, 167–182. doi: 10.1075/is.17031.wac

Zlatev, J. (2014). "The co-evolution of human intersubjectivity, morality, and language," in *The Social Origins of Language*, eds D. Dor, C. Knight, J. Lewis (Oxford: Oxford University Press), 249–266. doi: 10.1093/acprof:oso/9780199665327.003.0018

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

CONCEPTUAL ANALYSIS
published: 15 April 2021
doi: 10.3389/fpsyg.2021.651108

The Indexical Voice: Communication of Personal States and Traits in Humans and Other Primates

John L. Locke*

Lehman College, City University of New York, New York, NY, United Sates

Many studies of primate vocalization have been undertaken to improve our understanding of the evolution of language. Perhaps, for this reason, investigators have focused on calls that were thought to carry symbolic information about the environment. Here I suggest that even if these calls were in fact symbolic, there were independent reasons to question this approach in the first place. I begin by asking what kind of communication system would satisfy a species' biological needs. For example, where animals benefit from living in large groups, I ask how members would need to communicate to keep their groups from fragmenting. In this context, I discuss the role of social grooming and "close calls," including lip-smacking and grunting. Parallels exist in human societies, where information is exchanged about all kinds of things, often less about the nominal topic than the communicants themselves. This sort of indexical (or personal) information is vital to group living, which presupposes the ability to tolerate, relate to, and interact constructively with other individuals. Making indexical communication the focus of comparative research encourages consideration of somatic and behavioral cues that facilitate relationships and social benefits, including cooperation and collaboration. There is ample room here for a different and potentially more fruitful approach to communication in humans and other primates, one that focuses on personal appraisals, based on cues originating with individuals, rather than signals excited by environmental events.

Keywords: cues, signals, alarm calls, close calls, grooming, lip-smacking, small talk

OPEN ACCESS

Edited by:
D. Kimbrough Oller,
University of Memphis, United States

Reviewed by:
Drew Rendall,
University of New Brunswick
Fredericton, Canada
Adriano Lameira,
University of Warwick,
United Kingdom

****Correspondence:***
John L. Locke
john.locke@lehman.cuny.edu

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *08 January 2021*
Accepted: *24 February 2021*
Published: *15 April 2021*

Citation:
Locke JL (2021) The Indexical
Voice: Communication of Personal
States and Traits in Humans and
Other Primates.
Front. Psychol. 12:651108.
doi: 10.3389/fpsyg.2021.651108

INTRODUCTION

> "We do not really know what a man is saying until we know who he is and to whom he is speaking."
> F. G Bailey (1972), Gifts and Poison, 1972

In a book published in 1944, the physicist, Erwin Schrodinger, pointed out that the body of an organism survives by ingesting "negative entropy." Few knew what that meant, but as the psychologist, George Miller would later write, there are mathematical similarities between entropy and information. If the mind "survives by ingesting information," as he claimed, "all higher organisms are informavores" (Miller, 1983, p. 111).

Clearly, that designation includes humans. In the 1970s, sociologists began to use terms like "Information Age" and "Information Society" in reference to times and places where mediated communication systems were facilitating the flow of messages between individuals. But this information had an unusual property for humans and other mammals: neither the sender nor the receiver was physically present, nor did they necessarily know or care about each other. This arrangement may have reinforced a disembodied perspective on human communication.

Prior to the evolution of symbolic communication, humans would have regularly inferred each other's states and traits, purely from observation, as the other primates do. But in our species there was a tilt from individuals to messages, which may have diminished aspects of personal information, including the identity and nature of the participants. Here I suggest that this cultural effect may have caused us to think that humans and other primates are less alike than they actually are.

INFORMATION

For many decades, theories of human communication reflected the ideas of two signal engineers, Claude Shannon and Warren Weaver, who famously declared that if speakers tell listeners something they already know, no information passes between them (Shannon and Weaver, 1948). In this stripped-down view of communication, little thought was given to the possibility that speakers might communicate something besides the nominal topic, e.g., a willingness to share the material that they're expressing; a presumption that the listener doesn't know the thing that they're saying; or a belief that the listener would find it interesting.

Shannon and Weaver may have offered a reasonable interpretation of information theory, as formulated, but in real life people are often drawn to individuals with whom they share knowledge. It can be pleasurable, even exciting, to discover that friends know many of the same things that we do. Shared knowledge represents a form of inter-personal similarity which, like other instances of homophily, tends to promote affiliation (Launay and Dunbar, 2015).

In our lives, shared knowledge also lays the social and psychological groundwork for cooperation—an important issue to which we'll return later—but in a corporate context, working groups are often set up such that each member has something unique to contribute. The expectation is that team members will naturally pool their information, but the expectation is often unfulfilled due to the fact that a stronger disposition—to discuss common knowledge—gets in the way (Wittenbaum, 2000).

If the transmission of information is an important function of speech, then this should be revealed in various kinds of behavioral tests. For example, we might expect to find that novel information is more carefully articulated than familiar information, or is less guessable in the presence of noise, or—perhaps the best test—that speakers go out of their way to avoid saying anything that listeners might find ambiguous.

AMBIGUITY

If the transmission of information was the primary purpose of speaking, ambiguity would pose a serious threat to communication, but it's not clear that it does. Many of the things that people say are structurally ambiguous (e.g., Amy likes intelligent men and dogs), indicating that speakers rarely avoid this property, even when the possibilities for doing so are readily available (Ferreira and Dell, 2000; Ferreira et al., 2005; Haywood et al., 2005). Steven Pinker and his colleagues have pointed out that speakers not only tolerate ambiguity, they actively seek it. That's also true of indirect speech, e.g., where one diner asks another, "Can you reach the salt?" but gets the salt instead of an answer. These kinds of formulations are "inefficient, vulnerable to misunderstanding, and seemingly unnecessary" but are used universally (Pinker et al., 2008).

What I will emphasize here is that there are social factors that qualify, even mitigate, the value of semantic precision. In the first part of this paper, I venture into an empirical arena that is familiar to most primatologists, the transmission of information about environmental opportunities and dangers—specifically predators—raising questions about the relevance of this issue to its stated or implied context: the evolution of language. Then, I turn my attention to what I think of as "first principles," asking what kinds of communications would have been required in evolutionary history, given the social structure of humans and some primate species, and compare this to what actually exists. Here, I look at more promising areas of overlap among humans and other primates, specifically the socially functional but less semantically loaded communication registers related to "small talk" and, among the so-called "close calls," grunts and the lip-smacking associated with social grooming. Finally, I examine a wholly different class of information, one that pertains to the nature of the individual. I call this information "indexical," and suggest that human and non-human primates are remarkably similar when it comes to this class of inter-individual communication. Here I distinguish between information that is sent in the form of signals from information that is emitted in the form of cues. But first I will discuss something quite different, partly to make my point: the alarm calls of vervet monkeys.

ALARM CALLS

Some of the more exciting research on primate communication was done in East Africa on vervet monkeys, who were thought to issue a distinctive call in response to each of three classes of predators—snakes, eagles, and leopards. It was thought that these alarm calls warned other members of the group to take evasive action, and since vervet calls do not resemble the calls of the predators themselves, they were thought to be symbolic, much as words are.

If one were looking for informative uses of the voice in primates, these calls might seem to have qualified on several counts, and the initial reports were positive. The primary investigators, Robert Seyfarth and his colleagues, concluded that vervet monkeys "give different alarm calls to different predators. Recordings of the alarms played back when predators were absent caused the monkeys to run into trees for leopard alarms, look up for eagle alarms, and look down for snake alarms." (Seyfarth et al., 1980, p. 801). They even found that predator classification improved with age and experience.

The excitement was palpable. "Here," as several primatologists would later write, "there was evidence of language-like communication in a monkey, with the promise of similarly human-like cognitive complexity. The implications for language evolution were tantalizing." (Rendall and Owren, 2013, p. 153) If any primate vocalizations could be construed as meaningful, these calls would seem to be the closest thing to human speech that had been discovered to date. For, even if they were issued under conditions of extreme agitation and fear, alarm calls seemed to inform others in ways that appeared to be *verifiable*, surely a desirable criterion if these calls were certifiably to qualify as meaningful. But this attribute comes at a price, one that I will suggest is unacceptable. For, in search of signals whose meaning could be ascertained, researchers tended to avoid social vocalizations, which produce important but less discrete or observable responses (see descriptions in Silk, 2002; Silk et al., 2013).

If alarm calls were ambiguous, that is, if a call, like a yell, only meant that others should look out for *something,* I doubt that they would have excited much empirical attention, and in the end that seems to have been the result. Julia Fischer and her colleagues concluded that little in the way of supporting evidence was actually obtained or reported (Fischer et al., 2015; Fischer, 2017; also see Price et al., 2015). Moreover, vervets also gave similar sounding calls in other aggressive contexts.

There's irony here. While primatologists were trying to demonstrate that monkeys *mean something* with their calls, some linguists were demonstrating that many human utterances *mean little*—at least literally—given canned phrases such as *biting the hand that feeds you, in a nutshell, at the end of his rope, in the nick of time,* and *quit cold turkey* (Sidtis and Sidtis, 2018). So even if it was exciting to think that vervet monkeys *might* be using distinctive vocalizations to inform each other, there were issues to be considered if language was to be considered creative or generative as well as informative.

BUT, WHAT IF IT WERE TRUE?

The hypothesis of Seyfarth and his colleagues was ultimately disconfirmed, but what if agitated and frightened monkeys had verifiably transmitted predator-specific information? How would a handful of innate signals by individuals who were agitated or frightened, issued more or less reflexively, help us to understand language, a complex, unlearned medium that is used flexibly and socially, sometimes in jest? Fischer (2021) has suggested that vervet alarm calls may be innate. If that's true, can we say that vervets *do* something in order to inform each other, or merely *experience* events by reacting in an audible way?

Receptive components of primate vocalizations might even be innately present in another class of primates: humans. Thirty years ago, a team in Finland asked naïve humans to classify the affective content of macaque vocalizations that had been recorded when the animals were in situations associated with aggression and fear, sexual excitement, dominance and several other emotions. The listeners were extremely accurate in classifying the affective qualities of these vocalizations that they had never heard before (Leinonen et al., 1991).

Before moving on, I should restate my reason for bringing up alarm calls. It was not to add my voice, superfluously, to the conclusion that they may not be truly symbolic. It's to make a different sort of claim altogether: that studies of shrieking in a state of extreme agitation were never an appropriate way to find common ground with speech or language—cognitively, neurologically, or socially—whether the shrieks were precipitated by anything in the environment or not.

FIRST PRINCIPLES

When approaching the communication system of any species, it helps to begin by considering first principles, that is, the reasons why members of that species would need to communicate *at all*; and what kinds of information, and in what form, its members would benefit from exchanging. What kind of communications would have improved the lot of our evolutionary ancestors? If they needed to harmonize their interests, and to trust and cooperate with each other—which now seems obvious—they presumably required a means of communicating that would enable them to achieve these benefits. What kind of communication system would facilitate these objectives?

There are two simple questions here: what do members of a particular species do *in order to* communicate, or that *happens to* communicate, and what, given things we know about members of that species, would they be expected to do? Let's begin by reviewing some social facts, ones that may have played a role in evolution. The first relates to the effect of group size on external vigilance and within-group attention. When primate groups were small, there was considerable risk of predation, and the voice was needed to warn other members of the group whenever predators were detected (Port et al., 2020). Our distant ancestors kept an eye on the periphery of their tiny camps, where predators lay waiting.

When groups enlarged, members spent less time *looking for* predators and more time *looking at* members of their own group, in search of individuals with whom they might cooperate and collaborate (Locke, 2005). Doing so would have been essential, for if larger groups increased competition for resources, as is widely assumed, members would have needed innate mechanisms or strategies to keep their groups from splitting up.

Work by Robin Dunbar suggests that evolutionary increases in the size of primate groups produced new levels of social complexity, challenging and ultimately enhancing the interpretive capacity of the social brain and the use of vocalization to service relationships (Dunbar, 1992, 1993, 1998, 2009; also see Gustison et al., 2012; Roberts and Roberts, 2019). These changes in neural and vocal complexity may have been as adaptive as alarm calls. For if enlargement of groups helped to foil or reduce predation, other adaptations would have been needed to ensure that groups *retained their membership*. One of these may have been tonic communication, a call-response tactic that, according to Wolfgang Schleidt (1973, 1977), helps to keep groups together. There may also have been pressures to ramp up the analytical and interpretive abilities of receivers.

CLOSE CALLS

Many animals live in stable social groups and their fitness, according to Joan Silk and her colleagues, "depends at least in part on the outcome of their interactions with other group members." (Silk et al., 2013, p. 213). The success of these interactions is affected, in part, by the animals' use of "close calls" (Harcourt et al., 1993). Several classes of close calls, including lip- or tongue-smacking, grunts, and girneys, have been identified. As their name implies, these sounds are used by familiar individuals at close range, and with significant social consequences.

Any boost in primate sociality presupposes neural commitments to a mode of communication that would facilitate the evaluation of individuals for a variety of short- and long-term relationships. In a study of socialization and vocal behavior, McComb and Semple (2005) analyzed reports on forty-two different primate species, finding strong relationships among the size of primate groups, the time devoted to grooming, and size of vocal repertoires. Based on these findings, McComb and Semple suggested that a greater number of different vocalizations may be needed for animals to navigate complex networks of social relationships in primate societies.

RELATING

Speech enables people to perform in a public way, a behavior that particularly appeals to adult males in their quest for power or status and mating opportunities (Locke, 2001, 2011). But is the primary purpose of speech to perform, to inform, or to relate? In *The Tongues of Man*, the English phonetician, John Rupert Firth, wrote at length about the organs of speaking, including the tongue, lips, and jaw, but then shifted his attention to the "organs of talking." These, Firth said, "are at least two *normally associated human beings*." (Firth, 1937, p. 152).

One senses that Firth was onto something with his use of the word, *talking*, for it—like chatting—implies something about the social applications of speech. But what did Firth mean by talking? He didn't say. Nor did Darwin when he made a similar reference. He had heard naturalists remark that social animals who habitually use their vocal organs "as a means of intercommunication, use them *on other occasions* much more freely than other animals." (Darwin, 1872, p. 84, italics mine) What he meant by "other occasions" is unclear but, as I will discuss shortly, primate vocalizations are not limited to calls, nor do they invariably convey information about the physical environment.

If we look at the way people express themselves when relating to friends, we are likely to see something that is grammatically-simple, colloquial, predictable, redundant, structurally incomplete, and semantically imprecise. In fact, as I suggested earlier, much of it is not all that linguistic. If one wanted to study recursion, which some linguists take to be the hallmark of grammar, it's not clear that everyday conversational speech would be the best place to find it (Hauser et al., 2002). But if we were interested in comparing humans and other primates on the tendency to relate, the most suitable behavior would probably be manual grooming.

SOCIAL GROOMING

Manual grooming—sorting through the fur of an animal in search of parasites—may appear to be a nutritive process since groomers consume the yield, which contains protein. But grooming is primarily a social process. It tends to work upwardly, lower ranking animals being more likely to groom those of higher rank than the converse, and it acts like a favor (Cheney, 1977; Cheney and Seyfarth, 1990). de Waal (1997) has reported that animals are more likely to share food with another animal if they had previously groomed him than if they had not done so.

Grooming figures prominently into the formation and maintenance of social and cooperative relationships, but animals have other things to do besides groom. If grooming is performed dyadically, animals may be unable to maintain a satisfactory number of relationships. Dunbar (1993) proposed that polyadic conversations, which allow access to several social partners simultaneously, evolved as a form of social grooming to circumvent this time constraint. In non-human primates, it appears that polyadic grooming enables animals to maintain weak social relationships with many partners (Girrard-Buttoz et al., 2020).

Dunbar also suggested a second mechanism for the expansion of social relationships, what has come to be known as grooming-at-a-distance or, since that's physically impossible, vocal grooming. Malgorzata Arlet and her colleagues compared the rate of contact call exchanges between the females in two captive groups of Japanese macaques. They found a positive relationship between the time devoted to grooming by two females and the frequency with which they exchanged calls. Their results were consistent with predictions of the social bonding hypothesis, which holds that vocal exchanges can be interpreted as grooming-at-a-distance (Arlet et al., 2015; also see Kulahci et al., 2015).

LIP-SMACKING AND GRUNTING

Nearly a century ago, English zoologist Solly Zuckerman observed a colony of baboons and noticed "rhythmical lip, tongue, and jaw movements that usually accompany *friendly advances* between two animals, and that continue throughout the process of grooming." (Zuckerman, 1932, italics mine).

Forty years passed before anyone discussed the movements that were actually involved in lip-smacking. Then, one primatologist commented that "the actual smacking noise appears to be made by the *tongue* breaking contact with the roof of the mouth and/or upper lip or row of teeth, rather than by the lips themselves parting." (Redican, 1975, italics mine). But, for some reason, this activity came to be known as *lip*-smacking. Later, others measured and commented on lip-smacking's physical characteristics (e.g., Ghazanfar et al., 2012; Pereira et al., 2020), but it was the friendly advances that made lip-smacking interesting from a social standpoint.

The same goes for grunting, which is often used to signal peaceful intentions (Silk, 2002). In a study of wild Guinea baboons, Lauriane Faraut and her colleagues found that when approaching baboons grunted, they were more likely to interact in an affiliative fashion and less likely to displace the partner (Faraut et al., 2019). One could include other such studies here, but the point is that primates learn things about each other when they grunt and groom. "We are reasonably certain," wrote Joan Silk, "that monkeys make use of *information* derived from their own interactions with other group members to regulate their social relationships." (Silk, 2002, p. 153, my italics).

If there are classes of primate vocalization that convey information and *mean*, in effect, that an animal intends to be friendly—and on the basis of these vocalizations animals are able to form and maintain relationships—we are surely entitled to ask what is meant when we humans greet others (Laver, 1975).

WHAT DO "WE" MEAN?

When we hear a primate smack his lips, we don't ask what the individual smacks *mean*. We're aware that when it comes to semantics, it's not this smack or that smack, it's the act of *smacking*. "The medium," we might say, reminiscent of the 1970s media guru, Marshall McLuhan, "is the message." Since Austin (1962) and, before him, Peirce (1878), it has been clear that the medium—material that exposes the speaker's intention, whether it is, e.g., to praise, accuse, or belittle—can be a more important feature of a conversation than any of the words that are used.

In Zuckerberg's work, the animals that were advancing in a friendly way were Hamadryas baboons, but he could almost as easily have described a reunion, in our own species, of two friends who have just reconnected after a period of separation (Laver, 1975). It is that sort of friendly interaction that led social anthropologist, Bronislaw Malinowski, to propose the term, "phatic communion," for a sense of connection achieved by familiar individuals when speaking. Phatic communion, he said "serves to establish *bonds of personal union* between people brought together by the mere need of companionship and does not serve any purpose of communicating ideas." He added that "It is *only in certain very special uses among a civilized community, and only in its highest uses that language is employed to frame and express thoughts*" (Malinowski, 1923, p. 316).

Malinowski's view was inspired by his work in a small-scale society, but it has been observed in analyses of ordinary speech in more progressive cultures. One example was supplied by an American couple that allowed themselves to be recorded while on holiday. An analysis of nearly two thousand messages spoken by the couple revealed that fully three-fourths of their utterances were comments that involved *no* facts or other concrete information (Soskin and John, 1963).

The search for information in primate calls was understandable, given the desire to see them as referential, but much of human *speech* doesn't "mean" much, word for word, compared to the fact that the *speaker* has chosen to verbalize and has done so in a friendly manner, which may mean that he intends no harm, would like to interact, and is open to friendship. That's a huge message, even if it might seem to be small.

"SMALL TALK"

A century ago, in a short essay called "Small-Talk," an English writer described a semantically empty type of speech that is undertaken "not for the sake of saying something, but for the sake of saying anything" (Friedlaender, 1922). Forty-five years later, ethologist Desmond Morris offered a name for the sort of social speech the writer described. In *The Naked Ape*, he referred to "the *meaningless*, polite chatter of social occasions, the "nice weather we are having" or "have you read any good books lately" form of talking." The purpose of this sort of chatter, he said, is "to reinforce the greeting smile and to maintain the social togetherness." It *"is not concerned with the exchange of important ideas or information*." (Morris, 1967, p. 204, italics mine). He called this "meaningless" chatter, "grooming talking."

Reminiscent of Morris, philosopher Charles Taylor asked his readers to imagine that they were traveling with him on a train that is moving through a southern country. At some point, he says to a fellow passenger, "Whew, it's hot." This, he recognizes, "doesn't tell you anything you didn't know; neither that it is hot, nor that I find it so. Both these facts were plain to you before. Nor were they beyond your power to formulate; you probably already had formulated them."

What, then, was accomplished by this exchange? What it did, Taylor said, was "to create a rapport between us, the kind of thing which comes about when we do what we call striking up a conversation. Previously I knew that you were hot, and you knew that I was hot, and I knew that you must know that I knew that. But now it is out there as a fact *between us."* (Taylor, 1985, p. 273, italics mine).

Whether Taylor's traveler knew it or not, he was laying the groundwork for something else: cooperation. Suppose, for example, he recognized at some point that he needed to leave his seat for a few minutes. If so, he might feel comfortable asking his seat-mate to keep an eye on his things while he was gone, something he might feel less comfortable doing if they had not yet "broken the ice."

What few seem to have recognized is what might be occurring *during* small talk that is so cognitively undemanding that it can be processed by listeners while evaluating the speaker, which may be the primary purpose of the interaction, not the transmission of verbal information. Quiet conversation grants them the proximity and time required for the evaluation of weak somatic cues. A meta-analytic study by Balliet (2010) suggests that the mere act of communication enhances cooperation, especially in large groups of people. Which is interesting in light of primate work indicating that lip-smacking facilitates cooperation in wild chimpanzees (Fedurek et al., 2015).

There's one final point to be added here, one that is no less important. When we hear a message that has little semantic content, it is likely that the message is not the information that the speaker intended to convey, but the speaker himself: what he is like at the moment and may continue to be like in the future.

INDEXICALITY

Information *is* conveyed by the traveler in Taylor's anecdote, but it is *about the traveler himself*. From it, we may guess that he is, in the present situation, bored, lonely, open to interaction, feeling sociable, and a great many other things; and we may also infer that he is generally a friendly person. Given the length of the journey, one or more of these things may have been exactly what he needed to convey and his fellow passenger needed to know. The most important thing group-living individuals can know about others is *who they are,* that is, which of various individuals they happen to be, and what physical and behavioral characteristics they happen to possess.

In quiet conversations among familiars, what do utterances mean? Obviously, it depends on the topic, or does it? I suggest that what they mean, in the broadest and simplest sense is THE SPEAKER, as he was before and during the interaction. I suggest that the most basic and useful information that is orally communicated by humans (irrespective of culture) and other primates is personal, including information about behavioral dispositions in the moment, ones linked to transient physiological and emotional states, and reactions to observers, and stable tendencies to aggress, relate, or cooperate.

Which brings me to my main point: what would our group-living ancestors have needed to learn about each other? I will use the term *indexical* to represent characteristics of individuals who are emitting or sending information that is about them, whether it is in the form of transient physiological activity or emotion, or the expression of relatively stable features including temperament and personality. Fifty years ago, English phonetician David Abercrombie (1967, p. 9) used that term in reference to variations in a person's speech that "come and go according to his physical or mental state" (also see Peirce, 1878). For Abercombie, examples of physical and emotional states included excitement and nervousness, which directly affect the operation of the vocal organs, therefore, the voice and speech of the individual, producing "affective indices." "When a person speaks," wrote another British phonetician, John Laver, "he reveals often very detailed indexical information about his personal characteristics of regional origin, social status, personality, age, sex, state of health, mood, and a good deal more." (Laver, p. 221).

These phoneticians seized upon unintended variations of voice and speech precisely because they are unintended, therefore, like a nuisance variable, not properly considered a property of the language that was being described. Not being semantically critical, unintended material is likely to carry the most information about the speaker.

Whether speaking or not, Rendall and Owren (2013) have pointed out that we all communicate by way of a *biological code* which is neither arbitrary nor, necessarily, learned. This code carries information "about relevant social or physical characteristics of signalers such as their age, sex, body size, individual identity, emotional state or physical condition." (Rendall and Owren, 2013, p. 162). What they chose not to discuss is their value in linguistic communication, which may be to amplify or even to negate the literal meaning of any words that are embedded in the same acoustic stream.

Research now indicates that human listeners—even those who have never studied primates—are able to discriminate individual monkeys from their coos and screams, naturally and without training, much as they discriminate between members of our own species from their speech, even if limited to isolated vowel sounds (Owren and Rendall, 2003).

Earlier, I asked if vervets provide others with information or were merely experiencing something in a way that excited vocal behaviors that are audible to nearby others. Years ago, sociologist Irving Goffman (1959) offered two simple labels for the kinds of information that people exchange in their interactions. Some of it, he wrote, is "given," by which he meant material, typically in the form of words and gestures that we send to others deliberately. It is often self-serving and therefore may be unreliable. The donors are consciously aware that they are donating something and may even be able to anticipate its effect on recipients.

To describe the other kind of information that people communicate, Goffman used the term "given off." This information becomes available to others merely because they happen to be close enough to absorb it. People in sensory range discover things about us whether we want them to or not. Some of the cues are expressed by glands that emit chemicals into the atmosphere, announcing changes in physiological and emotional states. Others are leaked earlier in life, often under genetic influence, inscribing on the face and body lasting messages about the occupant.

CUES AND SIGNALS

In the evolution of vocal communication, it is important to distinguish between *cues*, the information that is given off, and *signals*, the information that is given. A cue to some physical or behavioral feature that is informative may occur as an *emission*. Examples include a loud voice, which may imply health or physical strength (Sell et al., 2010). But cues that are emitted can evolve into signals that are sent if their reproductive value is actively displayed or exaggerated (Maynard Smith and Harper, 1995; Fitch and Reby, 2001).

The existence of these cues may have contributed to the development of appraisal mechanisms that enabled our ancestors to cooperate selectively with individuals that had something to offer. I will suggest here that some primate species evolved ways to interpret the cues to transient states and stable traits, enabling them to select suitable partners for cooperative relationships.

In the last 20 years, a great deal has been learned about the physical cues to various personal qualities in humans. While primatologists have wondered whether non-human primates were capable of symbolic behavior, social psychologists have been asking how much humans learn about each other from various cues, some auditory or visual, others, thanks to recent and ongoing research, olfactory. It is not clear, at the moment, what all or even most cues to personal qualities are in primates, though some in the acoustic, chemical, visual, and tactile modalities have been identified (Moreira et al., 2013).

States

There are two broad classes of information that are given, or given off, by the communicants. One is whatever the individuals are experiencing in the moment, given the situation in which they find themselves. Much of this information, naturally, is emotional and is properly regarded as "affective" (Rendall and Owren, 2013). The rest of it is physiological, including changing levels of stress or sexual readiness. Both kinds of information, emotional and physiological, qualify as transient states.

Traits

Traits include whatever physical cues enable others to identify the caller or speaker as the individual that he or she uniquely is and a second class of information that includes what these individuals are like, that is, how they are best described in terms of personality or temperament and, in humans, character. These relatively *stable traits* are unusually important, for they enable others to predict future behaviors, therefore to approach or avoid individuals for mating or other cooperative activities in the future.

If personal traits seem particularly relevant to humans, it is important to recognize that non-human primates also emit cues to a number of physical and behavioral characteristics, ones that, for example, predict dominance, aggressiveness, and other behavioral dispositions. Like humans, the other primates also have histories, reputations, and essential qualities that are associated with temperament and personality.

We all know that humans have personal qualities, including temperaments, that appear early in development and continue well into adult life (Tang et al., 2020). In our species, these enduring traits are particularly important, given the need to make social and reproductive choices that have long-term consequences. What is the equivalent, if there is one, in other primate species? Whether we use the word "personalities" or not, it is recognized that some chimpanzees are gregarious, others bold or aggressive; some are risk-takers, open to new experiences, others are more introverted or shy. Significantly, some primates are known to be socially tolerant, a prerequisite to life in large groups in general and to cooperation specifically (Hare et al., 2007; Cieri et al., 2014; also Melis et al., 2006).

In a meta-analysis, Freeman and Gosling (2010) discovered a set of personality variables that had been reported in several hundred studies of primates, mostly adult chimpanzees and macaques. The most frequently shared traits were fearfulness, dominance, and confidence or aggressiveness, along with irritability, sociability, playfulness, and activity. Patrick Tkaczynski and his colleagues have reported that multiple types of social behavior were repeatable over the long term—up to 19 years—in wild chimpanzees. They concluded that "chimpanzees living in natural ecological settings have relatively stable long-term social phenotypes over years that may be independent of life-history or reproductive strategies." Their results, they said, "add to the growing body of the literature suggesting consistent individual differences in social tendencies are more likely the rule rather than the exception in group-living animals" (Tkaczynski et al., 2020, p. 1).

EMERGING POINTS OF AGREEMENT AND CONTINUITY

Though it has passed largely unrecognized, points of possible agreement between human and non-human primates have been quietly developing where indexical attributes are concerned. While primate research was exposing a gulf between human language and primate calls, a separate body of evidence was steadily exposing similarities between these species. For the fact is, when primatologists were looking for the seeds of speech, therefore language, evolutionary psychologists were looking for—and finding—a number of evolutionary antecedents to human communication.

Indexical Vocalization

Humans communicate a great deal of information about their states and traits. Some travels vocally. For example, there is evidence that men with low-pitched voices have more testosterone than other men, and are thought by female listeners to be more dominant and attractive (Collins, 2000; Feinberg et al., 2005; Puts et al., 2006), especially when their voice is heard in a courtship or mating context (Apicella and Feinberg, 2009; Little et al., 2011). This preference is stronger when women are in the fertile phase of their ovulatory cycle when estrogen levels are unusually high (Puts, 2005; Feinberg et al., 2006).

The human voice may be a reproductive cue when it varies with sex hormone levels, but it can also be appropriated for use as a reproductive signal. Lower vocal pitch predicts the mating success of males (Apicella et al., 2007). It has been reported that men lower their pitch, and that women may raise theirs, in a contrived mating context (Puts et al., 2006; Fraccaro et al., 2011). But, not everything is vocal.

The Indexical Face

In a reincarnation of classic physiognomy, it's been found that some personal traits can be accurately inferred from the dimensions of one's face—especially in men. Its the predictive value of a ratio—between bizygomatic width, that is, the lateral distance between left and right cheekbones (the zygions), and upper face height, that is, the vertical distance between the upper lip and the superior surface of the eyes. The typical ratios for adult females and males are about 1.80 and 1.86, respectively (Carré and McCormick, 2008). This difference is thought to reflect the fact that in adolescence, when the face usually elongates, a surge of testosterone tends to suppress this lengthening process in males (Ursi et al., 1993; Verdonck et al., 1999; Bulygina et al., 2006; Lefevre et al., 2013). The result is a face that is wider for its height than the average female face, with relatively larger cheekbones (Weston et al., 2004).

In humans, facial width and direct measures of testosterone predict many of the same things, including aggressiveness, in humans and several other primate species. In an interesting parallel, Carmen Lefevre and her colleagues observed a relationship in capuchin monkeys between facial width and alpha status (a proxy measure of aggressiveness) and a related measure of personality, assertiveness (Lefevre et al., 2014).

CONCLUDING REMARKS

Research on vocal signaling, undertaken in an evolutionary-linguistic context, implies the existence of a huge gulf between humans and other primates. That gulf narrows when one looks at vocal and other cues that are emitted in a social context. That is, humans and other primates are not as different as they have seemed. This conclusion rests not only on similarity at the level of function in social contexts but on the level of shared neural resources.

In this article I have called attention to the meaning or, better, the significance of social vocalizations in humans and other primates, vocalizations that carry personal information. If there is a lesson regarding speech, it is, as F. G. Bailey said, to understand what people are saying we must know who they are and who they are addressing, but much the same is true in the other primates. To understand their actions, we must know things about them, including their rank and reputation, and what they are experiencing at the moment. Without taking full account of these things, we also cannot interpret their behaviors, or them.

This forces us to consider an important question. If we are truly interested in communication, why limit studies of primate vocalization to signals that are broadcast, possibly in response to some attribute of the environment, when, in the case of personal cues, we are confronted with an embarrassment of riches, many in the form of personal information that is critical to a number of different social choices, including collaborators.

Whatever information we humans *think* we're providing listeners when we speak, a great deal of what we say and do, while speaking, carries information about us, including our reactions to listeners. This sort of information enabled our group-living ancestors to relate, that is, to form relationships, remedy disputes, coordinate activities, and cooperate on important projects, and it continues to do these things today.

As important as information about the emotional and physiological states of others can be, knowing something about the personality and temperament of others enables prediction of their future behaviors, which are of vital significance to the formation of long-term relationships. When physical cues to states and traits are studied across the primate classes, we develop opportunities to witness inter-specific continuity between humans an other primates than is possible when studies are limited to referential functions.

Through cross-disciplinary research that has been carried out in the past two decades, it has become clear that humans communicate many of the things that others need to know about them, as do the other primates. It seems to be the right time for a program of comparative research.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The work reported here was inspired by a Focus Group Funded by the Konrad Lorenz Institute for Evolution and Cognition (KLI) Research in Klosterneuburg, Austria and conducted at the KLI in the summer of 2019.

REFERENCES

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Apicella, C. L., and Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proc. R. Soc. Lond. B* 276, 1077–1082. doi: 10.1098/rspb.2008.1542

Apicella, C. L., Feinberg, D. R., and Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett.* 3, 682–684. doi: 10.1098/rsbl.2007.0410

Arlet, M., Jubin, R., Masataka, N., and Lemasson, A. (2015). Grooming-at-distance by exchanging calls in non-human primates. *Biol. Lett.* 11:711.

Austin, J. L. (1962). *How to do Things with Words*. Oxford: Clarendon Press.

Bailey, F. G. (1972). *Gifts and Poison: The Politics of Reputation*. Oxford: Basil Blackwell.

Balliet, D. (2010). Communication and cooperation in social dilemmas a meta-analytic review. *J. Conflict Resolution* 54, 39–57. doi: 10.1177/0022002709352443

Bulygina, E., Mitteroecker, P., and Aiello, L. (2006). Ontogeny of facial dimorphism and patterns of individual development within one human population. *Am. J. Phys. Anthropol.* 131, 432–443. doi: 10.1002/ajpa.20317

Carré, J. M., and McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proc. R. Soc. B* 275, 2651–2656. doi: 10.1098/rspb.2008.0873

Cheney, D., and Seyfarth, R. M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. Chicago, IL: University of Chicago Press.

Cheney, D. L. (1977). The acquisition of rank and the development of reciprocal alliances among free-ranging immature baboons. *Behav. Ecol. Sociobiol.* 2, 303–318. doi: 10.1007/bf00299742

Cieri, R. L., Churchill, S. E., Franciscus, R. G., Tan, J., and Hare, B. (2014). Craniofacial feminization, social tolerance, and the origins of behavioral modernity. *Curr. Anthropol.* 55, 419–443. doi: 10.1086/677209

Collins, S. A. (2000). Men's voices and women's choices. *Anim. Behav.* 60, 773–780.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.

de Waal, F. B. M. (1997). The chimpanzee's service economy: food for grooming. *Evol. Hum. Behav.* 18, 375–386. doi: 10.1016/s1090-5138(97)00085-8

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *J. Hum. Evol.* 22, 469–493. doi: 10.1016/0047-2484(92)90081-j

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behav. Brain Sci.* 16, 681–694. doi: 10.1017/s0140525x00032325

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol.* 6, 178–190.

Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Ann. Hum. Biol.* 36, 562–572. doi: 10.1080/03014460902960289

Faraut, L., Siviter, H., Dal Pesco, F., and Fischer, J. (2019). How life in a tolerant society affects the usage of grunts: evidence from female and male Guinea baboons. *Anim. Behav.* 153, 83–93. doi: 10.1016/j.anbehav.2019.05.003

Fedurek, P., Slocombe, K. E., Hartel, J. A., and Zuberbuhler, K. (2015). Chimpanzee lip-smacking facilitates cooperative behavior. *Sci. Rep.* 5:13460.

Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Smith, M. J. L., Cornwell, E. R., et al. (2005). The voice and face of woman: one ornament that signals quality? *Evol. Hum. Behav.* 26, 398–408. doi: 10.1016/j.evolhumbehav.2005.04.001

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., and DeBruine, L. M. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Horm. Behav.* 49, 215–222. doi: 10.1016/j.yhbeh.2005.07.004

Ferreira, V. S., and Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cogn. Psychol.* 40, 296–340. doi: 10.1006/cogp.1999.0730

Ferreira, V. S., Slevc, L. R., and Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition* 96, 263–284. doi: 10.1016/j.cognition.2004.09.002

Firth, J. R. (1937). *The Tongues of Men*. London: Watts & Co.

Fischer, J. (2017). Primate production and the riddle of language acquisition. *Psychon. Bull. Rev.* 24, 72–78. doi: 10.3758/s13423-016-1076-8

Fischer, J. (2021). Primate vocal communication and the evolution of speech. *Curr. Direct. Psychol. Sci.* [Epub ahead of print].

Fischer, J., Wheeler, B. C., and Higham, J. P. (2015). Is there any evidence for vocal learning in chimpanzee food calls? *Curr. Biol.* 25, 1028–1029.

Fitch, W. T., and Reby, D. (2001). The descended larynx is not uniquely human. *Proc. R. Soc. B Biol. Sci.* 268, 1669–1675. doi: 10.1098/rspb.2001.1704

Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., et al. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *J. Evol. Psychol.* 9, 57–67. doi: 10.1556/jep.9.2011.33.1

Freeman, H. D., and Gosling, S. D. (2010). Personality in nonhumaGhn primates: a review and evaluation of past research. *Am. J. Primatol.* 72, 653–671. doi: 10.1002/ajp.20833

Friedlaender, V. H. (1922). *Small-Talk. Pied Piper's Street and Other Essays*. Bristol: J. W. Arrowsmith.

Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., and Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Curr. Biol.* 22, 1176–1182. doi: 10.1016/j.cub.2012.04.055

Girrard-Buttoz, C., Surbeck, M., Samuni, L., Boesch, C., Fruth, B., Crockford, C., et al. (2020). Variable use of polyadic grooming and its effect on access to social partners in wild chimpanzees and bonobos. *Anim. Behav.* 168, 211–224. doi: 10.1016/j.anbehav.2020.08.021

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York, NY: Anchor Books.

Gustison, M. L., le Roux, A., and Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philos. Trans. R. Soc. B* 367, 1847–1859. doi: 10.1098/rstb.2011.0218

Harcourt, A. H., Stewart, K. J., and Hauser, M. D. (1993). Functions of wild gorilla 'close' calls. I. Repertoire, context, and interspecific comparison. *Behaviour* 124, 89–122. doi: 10.1163/156853993x00524

Hare, B., Melis, A. P., Woods, V., Hastings, S., and Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Curr. Biol.* 17, 619–623. doi: 10.1016/j.cub.2007.02.040

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

Haywood, S. L., Pickering, M. J., and Branigan, H. P. (2005). Do speakers avoid ambiguities during dialogue?. *Psychol. Sci.* 16, 362–366. doi: 10.1111/j.0956-7976.2005.01541.x

Kulahci, I. G., Rubenstein, D. I., and Ghazanfar, A. A. (2015). Lemurs groom-at-a-distance through vocal networks. *Anim. Behav.* 110, 179–186. doi: 10.1016/j.anbehav.2015.09.016

Launay, J., and Dunbar, R. I. M. (2015). Playing with strangers: which shared traits attract us most to new people? *PLoS One* 10:e0129688. doi: 10.1371/journal.pone.0129688

Laver, J. (1975). "Communicative functions of phatic communion," in *Organization of Behavior in Face-to-Face Interaction*, eds A. Kendon, R. M. Harris, and M. R. Key (The Hague: Mouton).

Lefevre, C. E., Lewis, G. J., Perrett, D. I., and Penke, L. (2013). Telling facial metrics: facial width is associated with testosterone levels in men. *Evol. Hum. Behav.* 34, 273–279. doi: 10.1016/j.evolhumbehav.2013.03.005

Lefevre, C. E., Wilson, V. D., Wilson, A. D., Morton, F. B., Brosnan, S. F., Paukner, A., et al. (2014). Facial width-to-height ratio relates to alpha status and assertive personality in capuchin monkeys. *PLoS One* 9:e0093369. doi: 10.1371/journal.pone.0093369

Leinonen, L., Linnankoski, I., Laakso, M.-L., and Aulanko, R. (1991). Vocal communication between species: man and macaque. *Lang. Commun.* 11, 241–262. doi: 10.1016/0271-5309(91)90031-p

Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C., and Roberts, S. C. (2011). Human preference for masculinity differs according to con- text in faces, bodies, voices, and smell. *Behav. Ecol.* 22, 862–868. doi: 10.1093/beheco/arr061

Locke, J. L. (2001). Rank and relationships in the evolution of spoken language. *J. R. Anthropol. Instit.* 7, 37–50. doi: 10.1111/1467-9655.00049

Locke, J. L. (2005). "Looking for, looking at: social control, honest signals, and intimate experience in human evolution and history," in *Animal Communication Networks*, ed. P. K. McGregor (Cambridge: Cambridge University Press).

Locke, J. L. (2011). *Duels and Duets: Why Men and Women Speak so Differently*. Cambridge, MA: University of Cambridge Press.

Malinowski, B. (1923). "The problem of meaning in primitive languages," in *The Meaning of Meaning*, eds C. K. Ogden and I. A. Richards (London: Routledge and Kegan Paul).

Maynard Smith, J., and Harper, D. G. C. (1995). Animal signals: models and terminology. *J. Theor. Biol.* 177, 305–311. doi: 10.1006/jtbi.1995.0248

McComb, K., and Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biol. Lett.* 1, 381–385. doi: 10.1098/rsbl.2005.0366

Melis, A. P., Hare, B., and Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science* 311, 1297–1300. doi: 10.1126/science.1123007

Miller, G. A. (1983). "Informavores," in *The Study of Information: Interdisciplinary Messages*, eds F. Machlup and U. Mansfield (Hoboken, NJ: Wiley-Interscience).

Moreira, L. A. A., Marques, E., Pessoa, A., and Cordeiro de Sousa, M. B. (2013). Socio-sexual communication: a review of the sensory modalities used by non-human primates. *Oecol. Austral.* 17, 113–129.

Morris, D. (1967). *The Naked Ape: A Zoologist's Study of the Human Animal*. New York: McGraw-Hill.

Owren, M. J., and Rendall, D. (2003). Salience of caller identity in Rhesus Monkey (*Macaca mulatta*) coos and screams: perceptual experiments with Human (Homo sapiens) listeners. *J. Comp. Psychol.* 117, 380–390. doi: 10.1037/0735-7036.117.4.380

Peirce, C. S. (1878). "How to make our ideas clear. Popular science monthly, 12, 286-302," in *A Chronological*, ed. C. S. Peirce (Bloomington, IN: Indiana University Press), 1872–1878.

Pereira, A. S., Kavanagh, E., Hobaiter, C., Slocombe, K. E., and Lameira, A. R. (2020). Chimpanzee lip-smacks confirm primate continuity for speech-rhythm evolution. *Biol. Lett.* 16:20200232. doi: 10.1098/rsbl.2020.0232

Pinker, S., Nowak, M. A., and Lee, J. J. (2008). The logic of indirect speech. *Proc. Natl. Acad. Sci. U.S.A.* 105, 833–838.

Port, M., Hildenbrandt, H. J., Pen, I., Schulke, O., Ostner, J., and Weissing, F. J. (2020). The evolution of social philopatry in female primates. *Am. J. Phys. Anthropol.* 173, 397–410. doi: 10.1002/ajp.24123

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5:13220.

Puts, D. A. (2005). Mating context and menstrual cycle phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* 26, 388–397. doi: 10.1016/j.evolhumbehav.2005.03.001

Puts, D. A., Gaulin, S. J. C., and Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* 27, 283–296. doi: 10.1016/j.evolhumbehav.2005.11.003

Redican, W. K. (1975). "Facial expressions in nonhuman primates," in *Primate Behavior: Developments in Field and Laboratory Research*, ed. L. A. Rosenblum (New York: Academic Press), 104–194.

Rendall, D., and Owren, M. J. (2013). "Communication without meaning or information: abandoning language-based and informational constructs in animal communication theory," in *Animal Communication Theory: Information and Influence*, ed. U. Stegmann (Cambridge: Cambridge University Press).

Roberts, A. I., and Roberts, S. G. B. (2019). Communicative roots of complex sociality and cognition. *Biol. Rev.* 95, 51-73.

Schleidt, W. M. (1973). Tonic communication: continual effects of discrete signs in animal communication systems. *J. Theor. Biol.* 42, 359–386. doi: 10.1016/0022-5193(73)90095-7

Schleidt, W. M. (1977). Tonic properties of animal communication systems. Part II. Tonic sources of tonic influences. *Ann. N. Y. Acad. Sci.* 290, 43–45. doi: 10.1111/j.1749-6632.1977.tb39715.x

Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., et al. (2010). Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. Lond. B* 277, 3509–3518. doi: 10.1098/rspb.2010.0769

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Shannon, C. E., and Weaver, W. (1948). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Sidtis, D. V. L., and Sidtis, J. J. (2018). The affective nature of formulaic language: a right-hemisphere subcortical process. *Front. Neurol.* 9:573. doi: 10.3389/fneur.2018.00573

Silk, J. (2002). "Grunts, girneys, and good intentions: the origins of strategic commitment in nonhuman primates," in *Commitment: Evolutionary perspectives*, ed. R. Nesse (New York: Russell Sage Press).

Silk, J., Cheney, D., and Seyfarth, R. (2013). A practical guide to the study of social relationships. *Evol. Anthropol.* 22, 213–225. doi: 10.1002/evan.21367

Soskin, W. F., and John, V. P. (1963). "The study of spontaneous talk," in *The Stream of Behaviour*, ed. R. G. Barker (New York: Springer).

Tang, A., Crawford, H., Morales, S., Degnan, K. A., Pine, D. S., and Fox, N. A. (2020). Infant behavioral inhibition predicts personality and social outcomes three decades later. *Proc. Natl. Acad. Sci. U.S.A.* 117, 9800–9807. doi: 10.1073/pnas.1917376117

Taylor, C. (1985). "The person," in *The Category of the Person: Anthropology, Philosophy, History*, eds M. Carrithers, S. Collins, and S. Lukes (Cambridge: Cambridge University Press).

Tkaczynski, P. J., Mielke, A., Samuni, L., Preis, A., and Wittig, R. M. (2020). Long-term repeatability in social behaviour suggests stable social phenotypes in wild chimpanzees. *R. Soc. Open Sci.* 7:945857.

Ursi, W. J. S., Trotman, C.-A., McNamara, J. A., and Behrents, R. G. (1993). Sexual dimorphism in normal craniofacial growth. *Angle Orthodont.* 63, 47–56.

Verdonck, A., Gaethofs, M., Carels, C., and deZegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *Eur. J. Orthod.* 21, 137–143. doi: 10.1093/ejo/21.2.137

Weston, E. M., Friday, A. E., Johnstone, R. A., and Schrenk, F. (2004). Wide faces or large canines? The attractive versus the aggressive primate. *Proc. R. Soc. Lond. B* 271, S416–S419.

Wittenbaum, G. M. (2000). The bias toward discussing shared information. *Commun. Res.* 27, 379–401. doi: 10.1177/009365000027003005

Zuckerman, S. (1932). *The Social Life of Monkeys and Apes*. New York: Harcourt, Brace, & Company.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 03 June 2021
doi: 10.3389/fpsyg.2021.648497

The Complexity and Phylogenetic Continuity of Laughter and Smiles in Hominids

Marina Davila-Ross[1*†] *and Guillaume Dezecache*[2†]

[1] *Psychology Department, University of Portsmouth, Portsmouth, United Kingdom,* [2] *Université Clermont Auvergne, CNRS, LAPSCO, Clermont-Ferrand, France*

Laughter and smiles are often, but not always, associated with positive affect. These expressions of humans help to promote social relationships as well as the development of cognitive and socio-emotional skills and they may have a positive impact on health and well-being, hereby covering a selection of fitness-relevant benefits. Both laughter and smiles of positive affect also occur early in human development and across cultures, suggesting deep roots in human biology. The present work provides an evolutionary reconstruction of the evolution of human laughter and smiles of positive affect in form and function, based on the principle of maximum parsimony. According to the Complexity and Continuity Hypothesis, human laughter and smiles of positive affect must have evolved within the context of play from ancestral species. Furthermore, ancestral ape laughter and their open-mouth faces must already have been complex in form and function and changed over time via categorically different phylogenetic pathways to become characteristic, effective, and pervasive behaviors of everyday social interactions in humans.

Keywords: laughter, smiles, laughing faces, open-mouth faces, evolution, principle of maximum parsimony

Edited by:
Irene M. Pepperberg,
Harvard University, United States

Reviewed by:
Nathalie Gontier,
University of Lisbon, Portugal
Katja Liebal,
Leipzig University, Germany

***Correspondence:**
Marina Davila-Ross
marina.davila-ross@port.ac.uk

†ORCID:
Marina Davila-Ross
orcid.org/0000-0002-5768-0217
Guillaume Dezecache
orcid.org/0000-0002-9366-6287

Specialty section:
This article was submitted to Evolutionary Psychology, a section of the journal Frontiers in Psychology

Received: *31 December 2020*
Accepted: *06 May 2021*
Published: *03 June 2021*

Citation:
Davila-Ross M and Dezecache G (2021) The Complexity and Phylogenetic Continuity of Laughter and Smiles in Hominids. Front. Psychol. 12:648497. doi: 10.3389/fpsyg.2021.648497

INTRODUCTION

Laughter and smiles of humans have often been discussed in close association with each other. Both play an important role in a wide range of daily social interactions (Owren and Bachorowski, 2003; Dezecache and Dunbar, 2012). They promote social cohesion (Provine, 2000; Dunbar and Mehu, 2008) as well as the development of cognitive and socio-emotional skills (Fredrickson, 2001; Gervais and Wilson, 2005) and they may even affect a person's health and well-being (Keltner and Bonanno, 1997; cf. Martin, 2002; Dunbar et al., 2012), all in all covering a selection of fitness-relevant benefits. Although it is important to be generally cautious when identifying emotional states of individuals based on their behavioral actions (Fridlund and Russell, 2006; Fridlund, 2014; Waller et al., 2017), laughter and smiles are clearly strongly linked to positive emotions in many situations and their corresponding neurochemical changes (Wild et al., 2003; Dunbar et al., 2012; Manninen et al., 2017). It is perhaps best observable as outbursts of affect in solitary contexts and in young children's play. Whereas, laughter and smiles often represent behavioral indicators of positive emotions in humans, which may serve in multiple ways, they can also be products of other emotions as well as functions (e.g., fear grins: van Hooff, 1972; Schadenfreude laughter: Szameitat et al., 2009a; for critical discussions on expressions and their relationships to emotions and communication, see Fridlund and Russell, 2006; Dezecache et al., 2013).

These expressions of positive affect seem to be deeply grounded in human biology. They occur in the first months of human development and continue to stay then mainly within positive contexts (Sroufe and Wunsch, 1972; Nwokah et al., 1994; Oller et al., 2013). Typically sounding laughter is even produced by those with no or close to no auditory experience, such as in congenitally deaf college students (Makagon et al., 2008). Furthermore, laughter and smiles occur in positive contexts across cultures, for instance in rough-and-tumble play of children (Grammer and Eibl-Eibesfeldt, 1990; Provine, 2000), and they are overall detected as positive expressions, albeit with cross-cultural differences in how they are more specifically processed (Masuda et al., 2008; Sauter et al., 2010; Bryant et al., 2016). Consequently, these positive expressions might reflect the more rudimentary, evolutionarily older forms of laughter and smiles and require special attention in the search for potential homologs in non-human primates ("primates" from here onwards).

It has long been noted that human and primate expressions of emotions may be similar in both anatomy and context, especially with regard to human laughter and smiles and the playful situations in which they may occur (Darwin, 1872; Andrew, 1963; Chevalier-Skolnikoff, 1973; Redican, 1982; Preuschoft, 1992). For instance, chimpanzee mothers may tickle their infants, who then produce play vocalizations and open-mouth faces (play faces), expressions they would also show during solitary play as well as play with peers, such as rough and tumble, tug of war, or play chase. Play vocalizations and open-mouth faces can be found among primates early in their development (Tomonaga et al., 2004; Bard et al., 2014) and across their different populations (e.g., in chimpanzees: Matsusaka, 2004; Davila-Ross et al., 2011). Interestingly, the way playing great apes produce their multimodal and unimodal expressions of play strongly resembles the way playing children produce laughter and open-mouth smiles of positive affect (Rothbart, 1973; Addyman et al., 2018), respectively. Whereas, such basic observations might naturally lead to the notion of phylogenetic continuity from primordial play expressions to human laughter and smiles of positive affect (Darwin, 1872; Redican, 1982), other possible explanations are that laughter and smiles are human-unique behaviors or that they evolved from different primordial expressions (van Hooff, 1972; Preuschoft and van Hooff, 1995).

In the last two decades, numerous in-depth studies on the form and function of primate play expressions were conducted that urge us to revisit the evolution of laughter and smiles. The goal of the present work is, thus, to examine these findings in combination with pioneering works on this topic in order to develop an evolutionary model of laughter and smiles, situated within the phylogeny of great apes and humans. Our evolutionary reconstruction from ancestral apes toward humans is mainly based on predictions we can make about the last common ancestor of extant great apes and humans, a relationship extracted from a multiplex phylogenetic clade that also includes other extinct species, which existed prior to the origin of modern humans. Because laughter is a multimodal expression that is primarily defined by its vocalization (Cosentino et al., 2016), we distinguish, when necessary in this work, between the terms "laugh vocalization" and "laughing face" to refer to its vocal and facial components, respectively (Ruch and Ekman, 2001; Drack et al., 2009). Smiles, in contrast, are the facial expressions that are not produced together with laugh vocalizations (Ekman et al., 1990; Iwase et al., 2002). Our evolutionary reconstruction is based on the principle of maximum parsimony. According to this principle, the most likely of alternative explanations on evolutionary pathways should involve the least number of predicted evolutionary steps for a given set of data (Saitou and Imanishi, 1989). It can be applied for any hard-wired multivariate traits (for primate expressions, see Geissmann, 2002; Davila-Ross and Geissmann, 2007; Davila-Ross et al., 2009).

DISCUSSION

Play Vocalizations and Laughter

The play vocalizations of great apes, among the primates, have received much research attention because of the acoustic similarity with human laughter found in the closest evolutionary relatives of humans (Darwin, 1872; van Lawick-Goodall, 1968; Gervais and Wilson, 2005; Leavens, 2009). These ape vocalizations are often, but not always, occurring as a series of low-frequency staccato grunts, which can perhaps most readily be elicited by tickling (Vettin and Todt, 2005; Davila-Ross and Zimmermann, 2009; Provine, 2017). With the aim to test for such potential homologies and to situate the evolution of laugh vocalizations within the larger phylogenetic trajectory of the Hominidae, Davila-Ross et al. (2009, 2010) used raw acoustic data obtained from tickling-induced vocalizations of infant and juvenile great apes and human infants to conduct phylogenetic analyses. Their generated maximum-parsimony trees matched the phylogeny of extant great apes and humans that has been well-established by geneticists (Ruvolo et al., 1994; Wildman et al., 2002; McBrearty and Jablonski, 2005). This match and additional analyses, that revealed robustness of the tree topology, indicated a shared evolutionary origin (Davila-Ross et al., 2009). Because human infant laughter was included in this study, phylogenetic evidence was provided that human laughter evolved from ancestral apes within the context of play at least 10–16 million years ago (Davila-Ross et al., 2009, 2010).

These laugh vocalizations of great apes occur predominantly during their dyadic play (Davila-Ross and Zimmermann, 2009), where they seem to help prolong such playful encounters in chimpanzees (Matsusaka, 2004; Davila-Ross et al., 2011). Even when recorded laughter of conspecifics was played back in two previous studies, chimpanzees did not produce laughter outside of play (Berntson et al., 1989; Davila-Ross et al., 2014), suggesting a limitation in flexible use, in contrast to human laughter (see Provine, 1992). Pilot video playback tests involving orangutans, gorillas and chimpanzees provided no different results (Davila-Ross, personal observations).

Within the context of play, however, there seem to be notable differences in laughter among these primates that suggest a higher level of complexity in both form and function in the African great apes. Whereas, orangutans (the great apes evolutionarily most distanced from humans) produce spontaneous laughter during their playful encounters, it is rare (Davila-Ross and Zimmermann, 2009). Instead, they often emit play squeaks, another type of play vocalization (Davila-Ross et al., 2010). Compared to their Asian counterparts, African great apes (gorillas, chimpanzees, and bonobos) laugh frequently during social play (Vettin and Todt, 2005; Davila-Ross and Zimmermann, 2009), suggesting a higher level of signal relevance. In addition, chimpanzees may emit laugh responses to their playmates' laughter that are shorter than their spontaneously produced laughter (Davila-Ross et al., 2011).

Furthermore, orangutans produce laughter of an overall simpler spectral and temporal structure than the African apes. Orangutan laugh bouts are typically short in duration, with uniformly noisy calls of mostly consistently egressive airflow (Davila-Ross et al., 2009). The African apes, in contrast, produce laugh bouts that are longer in duration, with more calls that are produced quickly and with more vibration regime changes (Davila-Ross et al., 2009). They may laugh alternating between ingressive and egressive airflow (also described as play panting: Matsusaka, 2004) as well as with sustained, consistently egressive airflow (Davila-Ross et al., 2010), an ability that enables the continuous flow of speech in humans (Winkworth et al., 1995; MacLarnon and Hewitt, 1999). Some chimpanzees and bonobos were even heard to produce laugh bouts for minutes, which was possible via both airflow systems (Davila-Ross, personal observations).

Human and great ape laugh vocalizations seem to differ bioacoustically and perceptually primarily in regular voicing, airflow direction and vibration regimes (Davila-Ross et al., 2009). Voicing occurs when the vocal folds are vibrating with a high degree of regularity, leading to distinctive melodic sounds that mark human speech (Lieberman, 1975; cf. Owren et al., 1997). It is present in some human laugh episodes, for instance "Ha-ha" and "He-he" (Provine and Yong, 1991; Provine, 2000), but rarely in great apes (Vettin and Todt, 2005; Davila-Ross et al., 2009). Interestingly, unvoiced human laughter, which includes grunt-, snort- and song-like laugh episodes, is more common than voiced laughter (Bachorowski et al., 2001). Furthermore, human laughter shows primarily egressive airflow and a notably higher abundance of quickly produced vibration regimes than that of great apes, which contributes to their spectral complexity (Davila-Ross et al., 2009, 2010; for human laugh acoustics also see Bachorowski et al., 2001; Szameitat et al., 2009b).

Open-Mouth Faces, Laughing Faces, and Smiles

Open-mouth faces of primates often occur during solitary play as well as social play and play invitations (Chevalier-Skolnikoff, 1974; Flack et al., 2004; Petru et al., 2009). Like play faces of mammals, in general, these primate expressions seem to guide play activities among the playmates by prolonging play and avoiding escalations into fights (Bekoff, 1995; Waller and Dunbar, 2005; Davila-Ross et al., 2011; Mancini et al., 2013). They can be spontaneously produced behaviors as well as responses to open-mouth faces of their playmates, for instance via rapid facial mimicry (Davila-Ross et al., 2008, 2011; Mancini et al., 2013; Palagi et al., 2019b).

In great ape play, open-mouth faces may be produced with laugh vocalizations as well as without them. Furthermore, these facial expressions show morphological commonalities with human laughing faces (see **Figure 1**) and smiles. Primate coding approaches that are based on the Facial Action Coding System (FACS: Ekman et al., 2002) provide special insight (for OrangFACS: Caeiro et al., 2013; for ChimpFACS: Vick et al., 2007). Such non-invasive methodologies allow researchers to systematically measure single facial movements of the underlying musculature shared by primates and humans to test for homologies (Ekman et al., 2002; Vick et al., 2007). Specifically, the open-mouth faces of play are marked by the contraction of the muscle *zygomaticus major*, which pulls the corners of the lips back and upwards, as well as by the opening of the lips (Parr et al., 2007; Davila-Ross et al., 2015; Waller et al., 2015), facial movements that characterize both laughing faces and smiles of humans (Ekman et al., 1990; Ruch and Ekman, 2001; Drack et al., 2009).

To test if human laughing faces and smiles emerged from pre-existing traits, Davila-Ross et al. (2015) measured with ChimpFACS the range of facial movements present in laughing chimpanzees during spontaneous play. The study revealed that laughing chimpanzees part their lips and pull them back as well as upwards while dropping their jaws; often they open their mouths further by raising their upper lips and sometimes they raise their cheeks, which causes crow's feet, besides showing other facial movements (Davila-Ross et al., 2015; see **Figure 1**). These movements matched those of laughing humans, which were identified by Drack et al. (2009) with FACS (also see Ekman et al., 1990; Ruch, 1993; Ruch and Ekman, 2001; Shiota et al., 2003). Moreover, the examined open-mouth faces of laughing chimpanzees included an almost identical set of facial configurations as the open-mouth faces without laughter (Davila-Ross et al., 2015), suggesting that they represent the same facial expressions. Therefore, based on the principle of maximum parsimony, the primordial open-mouth face of play represents the strongest candidate for a precursor of human laughing faces and smiles of positive affect—a phylogenetic reconstruction that involves no major evolutionary changes.

Indeed, the open-mouth face is the only primate expression that matches human laughing faces as well as human smiles in their close relationship to laughter, in addition to morphology and context. A different evolutionary model was provided by van Hooff (1972) in perhaps the most influential work on the evolution of human smiles over the past 50 years (see Lockard et al., 1977; Goldenthal et al., 1981; Preuschoft, 1992; Laidre and Yorzinski, 2005; Mehu and Dunbar, 2008), where he proposed another primate facial expression as the precursor of human smiles of positive affect, i.e., silent bared-teeth display (see **Figure 1**). According to van Hooff's (1972) Emancipation Hypothesis, the open-mouth variant of the silent bared-teeth display must have crossed contexts (from submissive to playful contexts) and morphologically converged with another facial display, i.e., open-mouth faces, as well as laugh vocalizations at a period following the last common ancestor of chimpanzees and humans in order to become smiles, emancipating then in function and crossing behavioral contexts in humans—a

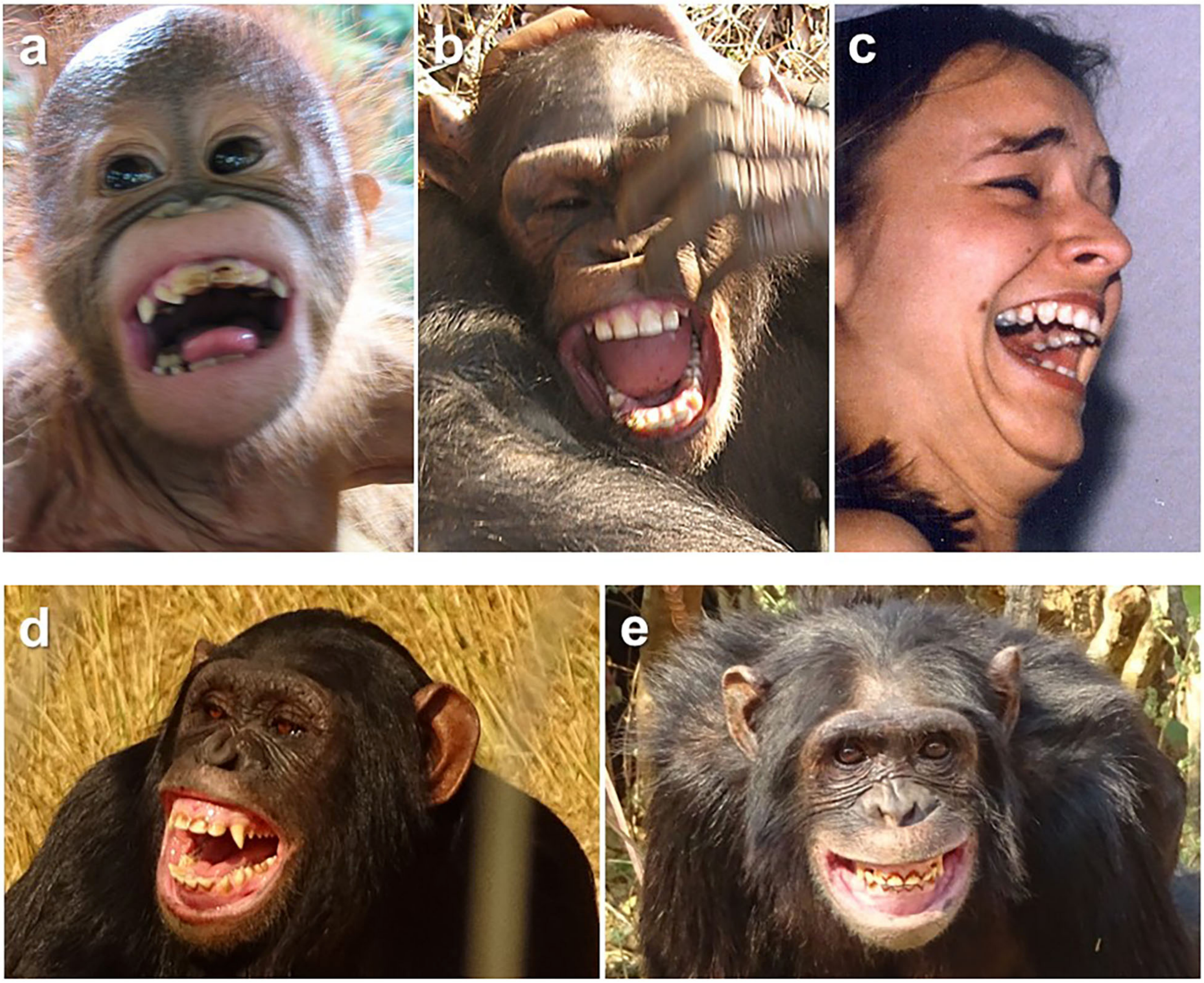


FIGURE 1 | Primate and human facial expressions: **(a)** orangutan and **(b)** chimpanzee open-mouth faces and **(c)** human laughing face; **(d)** open-mouth and **(e)** closed-mouth silent bared-teeth displays of chimpanzees (two silent bared-teeth display pictures, credit: Helene Chotard).

phylogenetic reconstruction which, like that of novel facial expressions, includes multiple major evolutionary changes.

It is important to note, however, that van Hooff (1972) also set the open-mouth face apart from other primate expressions as the most parsimonious explanation for a smile homolog, were it not for one then missing piece. Specifically, he pointed out that laughing children bare their upper tooth rows unlike the laughing chimpanzees he observed in Burgers' Zoo (van Hooff, 1972). Such open-mouth faces of chimpanzees and other primates are also referred to as relaxed open-mouth displays, a term that was used to indicate that during play the upper lip is relaxed, covering the upper teeth (van Hooff, 1972; Thierry et al., 1989). Over the last couple of decades, however, research demonstrated the exposure of upper teeth as part of the open-mouth face (also known as "relaxed open-mouth bared-teeth displays" and "full play faces") in laughing great apes (Davila-Ross and Zimmermann, 2009; Davila-Ross et al., 2015; see **Figure 1**) and in playing primates in general (van Hooff and Preuschoft, 2003; Palagi, 2006; Waller and Cherry, 2012). This facial movement is primarily caused by the *levator labii superioris* muscle contraction which raises the upper lip toward the nose. Therefore, the rationale for silent bared-teeth displays does not seem to hold any longer when it comes to smile precursors of positive affect.

Clearly the facial variants of open-mouth faces of play are closely linked to each other. In chimpanzee social play, approximately half of the 17 identified open-mouth configurations of play seem to involve an exposure of the upper teeth, configurations that may change into one another within a single display event (Davila-Ross et al., 2015; Davila-Ross, personal observations). The open-mouth variants marked by the exposed and covered upper teeth are, thus, to some extent interchangeable behaviors and they regularly occur in both gentle play and rough-and-tumble (see Davila-Ross and Zimmermann, 2009), although the upper teeth tend to be exposed more often during the latter play type (Palagi, 2006; Waller and Cherry, 2012). Whereas, primates, in general, make use of both open-mouth variants during play (e.g., geladas: Mancini et al., 2013; Japanese macaques: Scopa and Palagi, 2016: orangutans: Davila-Ross and Zimmermann, 2009; gorillas: Waller and Cherry, 2012; bonobos: de Waal, 1988), some primate species differ in the predominant use of these variants (see Thierry et al., 1989; van Hooff and Preuschoft, 2003; Scopa and Palagi, 2016). According to the Power Asymmetry Hypothesis by Preuschoft and van Hooff (1997), it should be particularly important for primate species living in strict linear dominance hierarchies (e.g., pig-tailed macaques) to produce distinct signals that can be easily recognized by their conspecifics, consequently shrinking the chances of escalations into fights, compared to primate species of a notably more relaxed social system (e.g., Tonkean macaques). Consistent with this hypothesis, pig-tailed macaques, for example, often do not expose their upper tooth rows during play, making their play signals distinct from silent bared-teeth displays of other contexts, much in contrast to Tonkean macaques (Bobbitt et al., 1964; Thierry et al., 1989; see van Hooff and Preuschoft, 2003).

Without the upper teeth exposed, open-mouth faces have only been infrequently documented outside of play (Preuschoft, 1992; Waller and Dunbar, 2005). Interestingly, Jan van Hooff's observed such an open-mouth face by a chimpanzee named Mama, which occurred after a known researcher revealed himself behind a leopard mask (van Hooff and Preuschoft, 2003). We are also aware of a youtube video clip where a juvenile orangutan produced this display following a magic trick (accessed 1st of April 2021: https://www.youtube.com/watch?v=OLrYzY3jVPY&ab_channel=Hydrasound). A similarly unusual incident took place at the Serengeti Park Hodenhagen, where a juvenile female chimpanzee named Pia was play inviting her father by pulling his hair, a clearly incongruent but seemingly harmless event (Davila-Ross, personal observation). As her father would not budge, Pia left and went to a different island of the enclosure, laid down on the grass and seemed to relax for a while, until she all of a sudden started producing open-mouth faces (Davila-Ross, personal observation; see **Supplementary Video**). It seems reasonable to conclude that Pia's outburst was induced by a representation of the preceding incongruent playful encounter. Collectively, such rare incidents provide evidence that great apes are able to produce open-mouth faces outside of play after non-aggressive violations of expectations, similar to human infant smiles (Reddy, 1991; for a discussion on benign violations and humor, see McGraw and Warren, 2010; Eckert et al., 2020). With the upper teeth exposed, open-mouth faces of play and the silent-bared teeth displays show interesting similarities that are discussed in the next section.

Complexity and Continuity Hypothesis of Laughter and Smiles

Empirical research on primate play expressions and human laughter and smiles of positive affect brings us back to the natural conclusion of phylogenetic continuity. Furthermore, with primordial open-mouth faces of play having evolved into human laughing faces and open-mouth smiles of positive affect, we can conclude that a shared ancestry of these two human facial expressions exists. Additional support comes from research on human facial morphology and physiology. Interestingly, human laughing faces and smiles of positive affect are both identified by *zygomaticus major* and include an overall similar configuration of facial muscle movements, such as *orbicularis oculi* muscle contractions (see Ekman et al., 1990; Shiota et al., 2003; Drack et al., 2009). Activations of *orbicularis oculi* raise the cheeks, causing the wrinkling around the eye corners, i.e., crow's feet, which characterizes Duchenne laughter/smiles, expressions arguably associated with felt positive emotions (Surakka and Hietanen, 1998; Ruch and Ekman, 2001; cf. Gunnery and Hall, 2015). In addition, a positron emission tomography (PET) scan study indicated that spontaneous laughter and smiles of positive affect, when produced by participants who watched funny videos, showed similar neural activations, predominantly in the bilateral supplementary motor area (SMA) and left putamen (Iwase et al., 2002).

Human laughter and smiles are also similar in function. Both may range from simple positive outbursts (Rothbart, 1973; Ekman et al., 1990) to highly complex behaviors, such as responses to humorous incidents and integral components in conversations (Owren and Bachorowski, 2003; Wild et al., 2003; Vettin and Todt, 2004; Arias et al., 2018). They promote cognitive and socio-emotional development (Fredrickson, 2001; Gervais and Wilson, 2005) and help to form, maintain and strengthen social relationships (Mehu et al., 2007; Dezecache and Dunbar, 2012; Wood and Niedenthal, 2018). Their positive effects may be further amplified when these expressions are shared among social partners (Provine, 1992; Hess and Bourgeois, 2010; cf. Dezecache et al., 2015) and when volitionally producing them (Bryant and Aktipis, 2014; Scott et al., 2014). Previously, human laughter and smiles have been identified as graded behaviors of intensity within positive contexts (Ekman, 1982; Redican, 1982; Bachorowski and Owren, 2001). According to the Diminutive Hypothesis, smiles have a lower arousal mode than laughter (Redican, 1982), a relationship that is also present in several languages (e.g., *rire* and *sourire* in French; *Lachen* and *Lächeln* in German). This hypothesis further implies that these two often interchangeable behaviors of positive contexts emerged from the same phylogenetic root (cf. Andrew, 1963; Redican, 1982), but it is hereby necessary to consider that laughter is predominantly a multimodal expression, unlike smiles. While human laugh faces are likely to be homologs of human smiles, empirical findings on primates suggest categorically different periods of evolutionary change for laugh vocalizations and smiles of positive affect. For laugh vocalizations, two main periods of evolutionary change among the hominids have been identified (Davila-Ross et al., 2010).

The first period of change for laugh vocalizations took place within great ape phylogeny. Laughter of the last common ancestor of extant great apes involved most likely a spontaneous, unvoiced vocalization of noise produced during social play, a vocalization that may have resembled a loud breathing (Davila-Ross et al., 2009). Laughter must have then become an increasingly complex and socially important vocalization, which was exchanged among playing conspecifics, as found in extant African apes (Davila-Ross et al., 2009). The second and predominant period of change for laugh vocalizations occurred closer to humans, after the divergence from a common ancestor with chimpanzees and bonobos. It must have been marked by regular voicing as well as consistently egressive airflow, two attributes of speech production (Davila-Ross et al., 2009). The increased presence of voicing in laughter may have heightened its level of perceived valence and arousal, as voiced laughter is processed as more positive in listeners than unvoiced laughter (Bachorowski and Owren, 2001). Interestingly, the main periods of change indicate categorically different phylogenetic pathways in laughter and smiles.

Specifically, open-mouth faces seem to have gone through one main period of evolutionary change in the past 10–16 million years. This facial expression of the last common ancestor of extant great apes was most likely already a behavior of high social relevance in play, with a complexity in both form and function, expressions that must have been used more flexibly than laughter (Davila-Ross and Zimmermann, 2009; Davila-Ross et al., 2015; Waller et al., 2015). Such complexity of open-mouth faces seems to be similarly present in monkeys (Mancini et al., 2013; Clark et al., 2020; see Preuschoft and van Hooff, 1995), perhaps even beyond primates (Palagi et al., 2019a; Taylor et al., 2019). Closer to humans and after the divergence from a common ancestor with chimpanzees and bonobos, these expressions may have involved more often *orbicularis oculi* muscle activations, possibly resulting in an increase in perceived valence and arousal (Ekman et al., 1990; Messinger et al., 2001; Soussignan, 2002; Davila-Ross et al., 2015), a change that is unrelated to the increase in voicing of laughter.

As part of the Complexity and Continuity Hypothesis, we therefore argue that great ape laughter and their open-mouth faces of play are homologs of the two arguably strongest behavioral indicators of positive affect in humans, expressions that are both frequently and, to some extent, similarly found in young children's play (Rothbart, 1973; Addyman et al., 2018). Consequently, humans are not unique in producing laughter and smiles of positively grounded motivations. Whereas, humans are known for having highly sophisticated social-cognitive abilities closely linked to cooperation unlike any other extant species (see Moll and Tomasello, 2007; Tomasello and Herrmann, 2010), these two important everyday expressions of social cohesion (Provine, 2000; Dunbar and Mehu, 2008) must have already existed on a pre-human basis, possibly to help initiate and prolong playful interactions with familiar conspecifics (Matsusaka, 2004; Waller and Dunbar, 2005; Davila-Ross et al., 2011; Mancini et al., 2013). Thus, the Complexity and Continuity Hypothesis contrasts the notion that human smiles of positive affect evolved within fear-related situations of ancestral species (see Emancipation Hypothesis: van Hooff, 1972; also see Andrew, 1963).

With the primate homologs of human laughter and smiles of positive affect identified based on the maximum parsimony principle, it is now important to critically evaluate how laughter and smiles became expressions of other motivations and functions (e.g., Schadenfreude: Szameitat et al., 2009a; mocking: Provine, 2000; embarrassed and polite smile: Ambadar et al., 2009). Regarding laugh vocalizations, it seems reasonable to conclude that such emancipation took place only after the last common ancestor of extant great apes and humans existed, because ape laughter is bioacoustically distinct (Davila-Ross et al., 2009; Taylor et al., 2021) and closely linked to play (Matsusaka, 2004; Davila-Ross et al., 2011), so that there cannot be variants with similar acoustic properties in other behavioral contexts. Furthermore, human infants within their first year of life produce various speech-related vocal types (protophones) free from contexts, but not laughter (Oller et al., 2013; for a comparative approach, see Dezecache et al., 2020). Consequently, laughter must have been used more flexibly closer toward humans, occurring in a wide range of everyday social interactions with gradually modifying acoustic properties (Owren and Bachorowski, 2003; Davila-Ross et al., 2010), perhaps accompanying key changes in language evolution (for laughter in conversation, see Vettin and Todt, 2004; Flamson and Bryant, 2013), when it also became an expression of other motivations and functions.

Regarding smiles of different motivations and functions, it is important to note that primates also expose their upper teeth when widely opening their mouths (e.g., silent-bared teeth displays and open-mouth threat faces) in contexts outside of play, namely in reconciliation, appeasement, affiliation, copulation and agonistic contexts (see Andrew, 1963; van Hooff, 1972; Weigel, 1979; Redican, 1982; Preuschoft, 1992; Liebal et al., 2004, 2006; Waller and Dunbar, 2005). As already mentioned in Jan van Hooff's (1972) important work on smile evolution, the silent bared-teeth displays show interesting similarities with open-mouth faces of play. More recently, FACS-based studies revealed that monkeys as well as apes may activate the same facial muscles across the two displays, although they seem to differ in their overall facial configurations (Parr and Waller, 2006; Parr et al., 2007; Davila-Ross et al., 2015; Waller et al., 2015; Clark et al., 2020) and perhaps in the motion pattern of muscle units, with the open-mouth faces of play showing the more dynamic pattern. Furthermore, these displays seem to instigate affiliative behaviors among interacting conspecifics across the behavioral contexts (Preuschoft, 1992; Bout and Thierry, 2005; Waller and Dunbar, 2005; de Marco and Visalberghi, 2007; Davila-Ross et al., 2011; Mancini et al., 2013).

Therefore, the possibility that open-mouth faces of play and silent bared-teeth displays are variants of each other should not be ignored, which means that the latter display might after all have had a role in smile evolution. In such case, however, its phylogenetic pathway would have been notably different from how it was presented in the Emancipation Hypothesis (see van Hooff, 1972; Preuschoft and van Hooff, 1995). Based on the data currently available, it is plausible that primordial forms of both open-mouth faces of play and silent bared-teeth displays emancipated in function, flexibly crossing behavioral contexts prior to the origin of hominids. This is further supported by data on the flexible use of primate facial expressions (Preuschoft, 1992; Waller and Dunbar, 2005; Davila-Ross et al., 2015; Scheider et al., 2016). An alternative explanation is that open-mouth faces of play are not related with silent bared-teeth displays, in which case the precursor of smiles of positive affect must have been used more freely across contexts after the last common ancestor of apes and humans existed, like laughter, to become a pervasive tool of human communication (see Owren and Bachorowski, 2003). More research is needed to test these two possible explanations. Both explanations, however, contrast with the Emancipation Hypothesis (van Hooff, 1972), where it was argued that fear-related displays emancipated in function closer toward human evolution.

In sum, the Complexity and Continuity Hypothesis of this work presents an evolutionary reconstruction of laughter and smiles of positive affect that reveals phylogenetic continuity. As evolution conserves hard-wired behavioral traits and their underlying processes rather than abolishes and rebuilds them, human laughter and smiles of positive affect must have evolved within the context of play in ancestral species. The Complexity and Continuity Hypothesis further states that their primordial displays must have already been complex in form and function ∼10–16 million years ago and further changed over time via categorically different phylogenetic pathways to become characteristic, effective and pervasive behaviors of everyday human social interactions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

MD-R and GD contributed to the write up. All authors contributed to the article and approved the submitted version.

FUNDING

The Royal Society Research Grant (RG140282) and Leakey Foundation Research Grant supported projects on laughter by MD-R. GD acknowledged the support received from the Agence Nationale de la Recherche of the French government through the program Investissements d'Avenir (16-IDEX-0001 CAP 20-25).

ACKNOWLEDGMENTS

We were thankful for very interesting previous discussions on the evolution of laughter and smiles with Elke Zimmermann, Michael Owren, Kim Bard, and Jan van Hooff.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.648497/full#supplementary-material

REFERENCES

Addyman, C., Fogelquist, C., Levakova, L., and Rees, S. (2018). Social facilitation of laughter and smiles in preschool children. *Front. Psychol.* 9:1048. doi: 10.3389/fpsyg.2018.01048

Ambadar, Z., Cohn, J. F., and Reed, L. I. (2009). All smiles are not created equal: morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *J. Nonverbal Behav.* 33, 17–34. doi: 10.1007/s10919-008-0059-5

Andrew, R. J. (1963). The origin and evolution of the calls and facial expressions of the primates. *Behaviour* 20, 1–107. doi: 10.1163/156853963X00220

Arias, P., Belin, P., and Aucouturier, J. J. (2018). Auditory smiles trigger unconscious facial imitation. *Curr. Biol.* 28, R782–R783. doi: 10.1016/j.cub.2018.05.084

Bachorowski, J. A., and Owren, M. J. (2001). Not all laughs are alike: voiced but not unvoiced laughter readily elicits positive affect. *Psychol. Sci.* 12, 252–257. doi: 10.1111/1467-9280.00346

Bachorowski, J. A., Smoski, M. J., and Owren, M. J. (2001). The acoustic features of human laughter. *J. Acoust. Soc. Am.* 110, 1581–1597. doi: 10.1121/1.1391244

Bard, K. A., Dunbar, S., Maguire-Herring, V., Veira, Y., Hayes, K. G., and McDonald, K. (2014). Gestures and social-emotional communicative development in chimpanzee infants. *Am. J. Primatol.* 76, 14–29. doi: 10.1002/ajp.22189

Bekoff, M. (1995). Play signals as punctuation: the structure of social play in canids. *Behaviour* 132, 419–429. doi: 10.1163/156853995X00649

Berntson, G. G., Boysen, S. T., Bauer, H. R., and Torello, M. S. (1989). Conspecific screams and laughter: cardiac and behavioral reactions of infant chimpanzees. *Dev. Psychobiol.* 22, 771–787. doi: 10.1002/dev.420220803

Bobbitt, R. A., Jensen, G. D., and Gordon, B. N. (1964). Behavioral elements (taxonomy) for observing mother-infant-peer interaction in *Macaca nemestrina*. *Primates* 5, 71–80. doi: 10.1007/BF01791645

Bout, N., and Thierry, B. (2005). Peaceful meaning for the silent bared-teeth displays of mandrills. *Int. J. Primatol.* 26, 1215–1228. doi: 10.1007/s10764-005-8850-1

Bryant, G. A., and Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evol. Hum. Behav.* 35, 327–335. doi: 10.1016/j.evolhumbehav.2014.03.003

Bryant, G. A., Fessler, D. M., Fusaroli, R., Clint, E., Aarøe, L., Apicella, C. L., et al. (2016). Detecting affiliation in colaughter across 24 societies. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4682–4687. doi: 10.1073/pnas.1524993113

Caeiro, C. C., Waller, B. M., Zimmermann, E., Burrows, A. M., and Davila-Ross, M. (2013). OrangFACS: a muscle-based facial movement coding system for orangutans (*Pongo* spp.). *Int. J. Primatol.* 34, 115–129. doi: 10.1007/s10764-012-9652-x

Chevalier-Skolnikoff, S. (1973). "Facial expression of emotion of nonhuman primates," in *Darwin and Facial Expression: A Century of Research in Review*, ed P. Ekman (New York, NY: Academic Press), 11–89.

Chevalier-Skolnikoff, S. (1974). The primate play face: a possible key to the determinants and evolution of play. *Rice Inst. Pamphlet Rice Univ. Stud.* 60, 9–29.

Clark, P. R., Waller, B. M., Burrows, A. M., Julle-Danière, E., Agil, M., Engelhardt, A., et al. (2020). Morphological variants of silent bared-teeth displays have different social interaction outcomes in crested macaques

(*Macaca nigra*). *Am. J. Phys. Anthropol*. 173, 411–422. doi: 10.1002/ajpa.24129

Cosentino, S., Sessa, S., and Takanishi, A. (2016). Quantitative laughter detection, measurement, and classification—a critical survey. *IEEE R. Bio Med. Eng*. 9, 148–162. doi: 10.1109/RBME.2016.2527638

Darwin, C. (1872). *The Expression of Emotion in Man and Animals*. London: Murray.

Davila-Ross, M., Allcock, B., Thomas, C., and Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion* 11:1013. doi: 10.1037/a0022594

Davila-Ross, M., and Geissmann, T. (2007). Call diversity of wild male orangutans: a phylogenetic approach. *Am. J. Primatol*. 69, 305–324. doi: 10.1002/ajp.20356

Davila-Ross, M., Hutchinson, J., Russell, J. L., Schaeffer, J., Billard, A., Hopkins, W. D., et al. (2014). Triggering social interactions: chimpanzees respond to imitation by a humanoid robot and request responses from it. *Anim. Cogn*. 17, 589–595. doi: 10.1007/s10071-013-0689-9

Davila-Ross, M., Jesus, G., Osborne, J., and Bard, K. A. (2015). Chimpanzees (*Pan troglodytes*) produce the same types of "laugh faces" when they emit laughter and when they are silent. *PLoS ONE* 10:e0127337. doi: 10.1371/journal.pone.0127337

Davila-Ross, M., Menzler, S., and Zimmermann, E. (2008). Rapid facial mimicry in orangutan play. *Biol. Lett*. 4, 27–30. doi: 10.1098/rsbl.2007.0535

Davila-Ross, M., Owren, M. J., and Zimmermann, E. (2009). Reconstructing the evolution of laughter in great apes and humans. *Curr. Biol*. 19, 1106–1111. doi: 10.1016/j.cub.2009.05.028

Davila-Ross, M., Owren, M. J., and Zimmermann, E. (2010). The evolution of laughter in great apes and humans. *Comm. Integrat. Biol*. 3, 191–194. doi: 10.4161/cib.3.2.10944

Davila-Ross, M., and Zimmermann, E. (2009). *Towards the Evolution of Laughter: A Comparative Analysis on Hominoids*. Saarbrücken: Suedwestdeutscher Verlag fuer Hochschulschriften.

de Marco, A., and Visalberghi, E. (2007). Facial displays in young tufted capuchin monkeys (*Cebus apella*): appearance, meaning, context, and target. *Folia Primatol*. 78, 118–137. doi: 10.1159/000097061

de Waal, F. B. (1988). The communicative repertoire of captive Bonobos (*Pan paniscus*) compared to that of chimpanzees. *Behaviour* 106, 183–251. doi: 10.1163/156853988X00269

Dezecache, G., and Dunbar, R. I. (2012). Sharing the joke: the size of natural laughter groups. *Evol. Hum. Behav*. 33, 775–779. doi: 10.1016/j.evolhumbehav.2012.07.002

Dezecache, G., Jacob, P., and Grezes, J. (2015). Emotional contagion: its scope and limits. *Trends Cogn. Sci*. 19, 297–299. doi: 10.1016/j.tics.2015.03.011

Dezecache, G., Mercier, H., and Scott-Phillips, T. C. (2013). An evolutionary approach to emotional communication. *J. Pragmat*. 59, 221–233. doi: 10.1016/j.pragma.2013.06.007

Dezecache, G., Zuberbühler, K., Davila-Ross, M., and Dahl, C. D. (2020). Flexibility in wild infant chimpanzee vocal behavior. *J. Lang. Evol*. 6:lzaa009. doi: 10.1093/jole/lzaa009

Drack, P., Huber, T., and Ruch, W. (2009). "The apex of happy laughter: a FACS-study with actors," in *Current and Future Perspectives in Facial Expression Research: Topics and Methodical Questions*, eds E. Bänninger-Huber and D. Peham (Innsbruck: Innsbruck University Press), 32–37.

Dunbar, R., and Mehu, M. (2008). Naturalistic observations of smiling and laughter in human group interactions. *Behaviour* 145, 1747–1780. doi: 10.1163/156853908786279619

Dunbar, R. I., Baron, R., Frangou, A., Pearce, E., van Leeuwen, E. J., Stow, J., et al. (2012). Social laughter is correlated with an elevated pain threshold. *Proc. Roy. Soc. B Biol. Sci*. 279, 1161–1167. doi: 10.1098/rspb.2011.1373

Eckert, J., Winkler, S. L., and Cartmill, E. A. (2020). Just kidding: the evolutionary roots of playful teasing. *Biol. Lett*. 16:20200370. doi: 10.1098/rsbl.2020.0370

Ekman, P. (1982). *Emotion in the Human Face*. New York, NY: Cambridge University Press.

Ekman, P., Davidson, R. J., and Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology. *J. Pers. Soc. Psychol*. 58, 342–353. doi: 10.1037/0022-3514.58.2.342

Ekman, P., Friesen, W. V., and Hager, J. C. (2002). *Facial Action Coding System*. Salt Lake City, UT: Research Nexus.

Flack, J. C., Jeannotte, L. A., and de Waal, F. (2004). Play signaling and the perception of social rules by juvenile chimpanzees (*Pan troglodytes*). *J. Comp. Psychol*. 118, 149–159. doi: 10.1037/0735-7036.118.2.149

Flamson, T. J., and Bryant, G. A. (2013). "Signals of humor: encryption and laughter in social interaction," in *Developments in Linguistic Humour Theory*, ed M. Dynel (Amsterdam: John Benjamins), 49–74.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *Am. Psychol*. 56, 218–226. doi: 10.1037/0003-066X.56.3.218

Fridlund, A. J. (2014). *Human Facial Expression: An Evolutionary View*. San Diego, CA: Academic Press.

Fridlund, A. J., and Russell, J. A. (2006). "The functions of facial expressions: what's in a face?" in *The Sage Handbook of Nonverbal Communication*, eds V. Manusov and M. L. Patterson (Thousand Oaks, CA: Sage Publications, Inc.), 299–319.

Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biol. Rev*. 77, 57–76. doi: 10.1017/S1464793101005826

Gervais, M., and Wilson, D. S. (2005). The evolution and functions of laughter and humor: a synthetic approach. *Q. Rev. Biol*. 80, 395–430. doi: 10.1086/498281

Goldenthal, P., Johnston, R. E., and Kraut, R. E. (1981). Smiling, appeasement, and the silent bared-teeth display. *Ethol. Sociobiol*. 2, 127–133. doi: 10.1016/0162-3095(81)90025-X

Grammer, K., and Eibl-Eibesfeldt, I. (1990). "The ritualisation of laughter," in *Natürlichkeit der Sprache und der Kultur*, ed W. A. Koch (Bochum: Universitätsverlag), 192–214.

Gunnery, S. D., and Hall, J. A. (2015). "The expression and perception of the Duchenne smile," in *The Social Psychology of Nonverbal Communication*, eds A. Kostic and D. Chadee (London: Palgrave Macmillan), 114–133.

Hess, U., and Bourgeois, P. (2010). You smile–I smile: emotion expression in social interaction. *Biol. Psychol*. 84, 514–520. doi: 10.1016/j.biopsycho.2009.11.001

Iwase, M., Ouchi, Y., Okada, H., Yokoyama, C., Nobezawa, S., Yoshikawa, E., et al. (2002). Neural substrates of human facial expression of pleasant emotion induced by comic films: a PET study. *Neuroimage* 17, 758–768. doi: 10.1006/nimg.2002.1225

Keltner, D., and Bonanno, G. A. (1997). A study of laughter and dissociation: distinct correlates of laughter and smiling during bereavement. *J. Pers. Soc. Psychol*. 73, 687–702. doi: 10.1037/0022-3514.73.4.687

Laidre, M. E., and Yorzinski, J. L. (2005). The silent bared-teeth face and the crest-raise of the mandrill (*Mandrillus sphinx*): a contextual analysis of signal function. *Ethology* 111, 143–157. doi: 10.1111/j.1439-0310.2004.01051.x

Leavens, D. A. (2009). Animal communication: laughter is the shortest distance between two apes. *Curr. Biol*. 19, R511–R513. doi: 10.1016/j.cub.2009.05.007

Liebal, K., Pika, S., and Tomasello, M. (2004). Social communication in siamangs (*Symphalangus syndactylus*): use of gestures and facial expressions. *Primates* 45, 41–57. doi: 10.1007/s10329-003-0063-7

Liebal, K., Pika, S., and Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6, 1–38. doi: 10.1075/gest.6.1.02lie

Lieberman, P. (1975). *On the Origins of Language: An Introduction to the Evolution of Human Speech*. New York, NY: Macmillan.

Lockard, J. S., Fahrenbruch, C. E., Smith, J. L., and Morgan, C. J. (1977). Smiling and laughter: different phyletic origins? *B. Psychonomic Soc*. 10, 183–186. doi: 10.3758/BF03329317

MacLarnon, A., and Hewitt, G. P. (1999). The evolution of human speech: the role of enhanced breathing control. *Am. J. Phys. Anthropol*. 109, 341–363.

Makagon, M. M., Funayama, E. S., and Owren, M. J. (2008). An acoustic analysis of laughter produced by congenitally deaf and normally hearing college students. *J. Acoust. Soc. Am*. 124, 472–483. doi: 10.1121/1.2932088

Mancini, G., Ferrari, P. F., and Palagi, E. (2013). In play we trust. Rapid facial mimicry predicts the duration of playful interactions in geladas. *PLoS ONE* 8:e66481. doi: 10.1371/journal.pone.0066481

Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., et al. (2017). Social laughter triggers endogenous opioid release in humans. *J. Neurosci*. 37, 6125–6131. doi: 10.1523/JNEUROSCI.0688-16.2017

Martin, R. A. (2002). Is laughter the best medicine? humor, laughter, and physical health. *Curr. Dir. Psychol. Sci*. 11, 216–220. doi: 10.1111/1467-8721.00204

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., and Van de Veerdonk, E. (2008). Placing the face in context: cultural differences in the perception of facial emotion. *J. Pers. Soc. Psychol*. 94, 365–381. doi: 10.1037/0022-3514.94.3.365

Matsusaka, T. (2004). When does play panting occur during social play in wild chimpanzees? *Primates* 45, 221–229. doi: 10.1007/s10329-004-0090-z

McBrearty, S., and Jablonski, N. G. (2005). First fossil chimpanzee. *Nature* 437, 105–108. doi: 10.1038/nature04008

McGraw, A. P., and Warren, C. (2010). Benign violations: making immoral behavior funny. *Psychol. Sci*. 21, 1141–1149. doi: 10.1177/0956797610376073

Mehu, M., and Dunbar, R. I. (2008). Relationship between smiling and laughter in humans (*Homo sapiens*): testing the power asymmetry hypothesis. *Folia Primatol*. 79, 269–280. doi: 10.1159/000126928

Mehu, M., Grammer, K., and Dunbar, R. I. (2007). Smiles when sharing. *Evol. Hum. Behav*. 28, 415–422. doi: 10.1016/j.evolhumbehav.2007.05.010

Messinger, D. S., Fogel, A., and Dickson, K. L. (2001). All smiles are positive, but some smiles are more positive than others. *Dev. Psychol*. 37, 642–653. doi: 10.1037/0012-1649.37.5.642

Moll, H., and Tomasello, M. (2007). Cooperation and human cognition: the Vygotskian intelligence hypothesis. *Philos. Trans. R. Soc. B* 362, 639–648. doi: 10.1098/rstb.2006.2000

Nwokah, E. E., Hsu, H.-C., Dobrowolska, O., and Fogel, A. (1994). The development of laughter in mother-infant communication: timing parameters and temporal sequences. *Infant Behav. Dev*. 17, 23–35. doi: 10.1016/0163-6383(94)90019-1

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., and Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6318–6323. doi: 10.1073/pnas.1300337110

Owren, M. J., and Bachorowski, J. A. (2003). Reconsidering the evolution of nonlinguistic communication: the case of laughter. *J. Nonverbal Behav*. 27, 183–200. doi: 10.1023/A:1025394015198

Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cynocephalus ursinus*): implications for production processes and functions. *J. Acoust. Soc. Am*. 101, 2951–2963. doi: 10.1121/1.418523

Palagi, E. (2006). Social play in bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*): implications for natural social systems and interindividual relationships. *Am. J. Phys. Anthropol*. 129, 418–426. doi: 10.1002/ajpa.20289

Palagi, E., Marchi, E., Cavicchio, P., and Bandoli, F. (2019b). Sharing playful mood: rapid facial mimicry in *Suricata suricatta*. *Anim. Cogn*. 22, 719–732. doi: 10.1007/s10071-019-01269-y

Palagi, E., Norscia, I., Pressi, S., and Cordoni, G. (2019a). Facial mimicry and play: a comparative study in chimpanzees and gorillas. *Emotion* 19, 665–681. doi: 10.1037/emo0000476

Parr, L. A., and Waller, B. M. (2006). Understanding chimpanzee facial expression: insights into the evolution of communication. *Soc. Cogn. Affect. Neurosci*. 1, 221–228. doi: 10.1093/scan/nsl031

Parr, L. A., Waller, B. M., Vick, S. J., and Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion* 7, 172–181. doi: 10.1037/1528-3542.7.1.172

Petru, M., Špinka, M., Charvátová, V., and Lhota, S. (2009). Revisiting play elements and self-handicapping in play: a comparative ethogram of five Old World monkey species. *J. Comp. Psychol*. 123, 250–263. doi: 10.1037/a0016217

Preuschoft, S. (1992). "Laughter" and "smile" in Barbary macaques (*Macaca sylvanus*). *Ethology* 91, 220–236. doi: 10.1111/j.1439-0310.1992.tb00864.x

Preuschoft, S., and van Hooff, J. A. R. A.M. (1995). Homologizing primate facial displays: a critical review of methods. *Folia Primatol*. 65, 121–137. doi: 10.1159/000156878

Preuschoft, S., and van Hooff, J. A. R. A. M. (1997). "The social function of "smile" and "laughter": variations across primate species and societies," in *Nonverbal Communication: Where Nature Meets Culture*, eds U. C. Segerstråle and P. Molnár (Mahwah, NJ: Lawrence Erlbaum Associates), 171–190.

Provine, R. R. (1992). Contagious laughter: laughter is a sufficient stimulus for laughs and smiles. *Bull. Psychonom. Soc*. 30, 1–4. doi: 10.3758/BF03330380

Provine, R. R. (2000). *Laughter: A Scientific Investigation*. New York, NY: Penguin Books.

Provine, R. R. (2017). Laughter as an approach to vocal evolution: the bipedal theory. *Psychon. B. Rev*. 24, 238–244. doi: 10.3758/s13423-016-1089-3

Provine, R. R., and Yong, Y. L. (1991). Laughter: a stereotyped human vocalization. *Ethology* 89, 115–124. doi: 10.1111/j.1439-0310.1991.tb00298.x

Reddy, V. (1991). "Playing with others' expectations: Teasing and mucking about in the first year," in *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading*, eds A. Whiten and R.W. Byrne (Oxford: Basil Blackwell), 143–158.

Redican, W. K. (1982). "An evolutionary perspective on human facial displays," in *Emotion in the Human Face*, ed P. Ekman (Cambridge: Cambridge University Press), 212–281.

Rothbart, M. K. (1973). Laughter in young children. *Psychol. Bull*. 80, 247–256. doi: 10.1037/h0034846

Ruch, W. (1993). "Exhilaration and Humor," in *The Handbook of Emotion*, eds M. Lewis and J.M. Haviland (New York, NY: Guilford Publications), 605–616.

Ruch, W., and Ekman, P. (2001). "The expressive pattern of laughter," in *Emotion, Qualia, and Consciousness*, ed A. Kaszniak (Tokyo: World Scientific), 426–443.

Ruvolo, M., Pan, D., Zehr, S., Goldberg, T., and Disotell, T. R. (1994). Gene trees and hominoid phylogeny. *Proc. Natl. Acad. Sci. U.S.A*. 91, 8900–8904. doi: 10.1073/pnas.91.19.8900

Saitou, N., and Imanishi, T. (1989). Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. *Mol. Biol. Evol*. 6, 514–525.

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2408–2412. doi: 10.1073/pnas.0908239106

Scheider, L., Waller, B. M., Oña, L., Burrows, A. M., and Liebal, K. (2016). Social use of facial expressions in hylobatids. *PLoS ONE* 11:e0151733. doi: 10.1371/journal.pone.0151733

Scopa, C., and Palagi, E. (2016). Mimic me while playing! Social tolerance and rapid facial mimicry in macaques (Macaca tonkeana and Macaca fuscata). *J. Comp. Psychol*. 130, 153–161. doi: 10.1037/com0000028

Scott, S. K., Lavan, N., Chen, S., and McGettigan, C. (2014). The social life of laughter. *Trends Cogn. Sci*. 18, 618–620. doi: 10.1016/j.tics.2014.09.002

Shiota, M. N., Campos, B., and Keltner, D. (2003). The faces of positive emotion: prototype displays of awe, amusement, and pride. *Ann. N. Y. Acad. Sci*. 1000, 296–299. doi: 10.1196/annals.1280.029

Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity: a test of the facial feedback hypothesis. *Emotion* 2, 52–74. doi: 10.1037/1528-3542.2.1.52

Sroufe, L. A., and Wunsch, J. P. (1972). The development of laughter in the first year of life. *Child Dev*. 43, 1326–1344. doi: 10.2307/1127519

Surakka, V., and Hietanen, J. K. (1998). Facial and emotional reactions to Duchenne and non-Duchenne smiles. *Int. J. Psychophysiol*. 29, 23–33. doi: 10.1016/S0167-8760(97)00088-3

Szameitat, D. P., Alter, K., Szameitat, A. J., Darwin, C. J., Wildgruber, D., Dietrich, S., et al. (2009a). Differentiation of emotions in laughter at the behavioral level. *Emotion* 9, 397–405. doi: 10.1037/a0015692

Szameitat, D. P., Alter, K., Szameitat, A. J., Wildgruber, D., Sterr, A., and Darwin, C. J. (2009b). Acoustic profiles of distinct emotional expressions in laughter. *J. Acoust. Soc. Am*. 126, 354–366. doi: 10.1121/1.3139899

Taylor, D., Dezecache, G., and Davila-Ross, M. (2021). Filling in the gaps: acoustic gradation increases in the vocal ontogeny of chimpanzees (*Pan troglodytes*). *Am. J. Primatol.* 83:e23249. doi: 10.1002/ajp.23249

Taylor, D., Hartmann, D., Dezecache, G., Te Wong, S., and Davila-Ross, M. (2019). Facial complexity in sun bears: exact facial mimicry and social sensitivity. *Sci. Rep*. 9, 1–6. doi: 10.1038/s41598-019-39932-6

Thierry, B., Demaria, C., Preuschoft, S., and Desportes, C. (1989). Structural convergence between silent bared-teeth display and relaxed open-mouth display in the Tonkean macaque (*Macaca tonkeana*). *Folia Primatol*. 52, 178–184. doi: 10.1159/000156396

Tomasello, M., and Herrmann, E. (2010). Ape and human cognition: what's the difference? *Curr. Dir. Psychol. Sci*. 19, 3–8. doi: 10.1177/0963721409359300

Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., Kosugi, D., Mizuno, Y., et al. (2004). Development of social cognition in infant chimpanzees (*Pan troglodytes*): face recognition, smiling, gaze, and the lack of triadic interactions 1. *Jpn. Psychol. Res*. 46, 227–235. doi: 10.1111/j.1468-5584.2004.00254.x van Hooff, J. A. R. A.M. (1972). "A comparative approach to the phylogeny of laughter and smiling," in *Non-verbal Communication*, ed R.A. Hinde (Cambridge: Cambridge University Press), 209–241.

van Hooff, J. A. R. A.M., and Preuschoft, S. (2003). "Laughter and smiling: the intertwining of nature and culture," in *Animal Social Complexity: Intelligence, Culture, and Individualized Societies*, eds F. B. M. de Waal and P. L. Tyack (Cambridge: Harvard University Press), 261–287.

van Lawick-Goodall, J. (1968). The behaviour of free-living chimpanzees in the Gombe Stream area. *Anim. Behav. Monogr*. 1, 65–311. doi: 10.1016/S0066-1856(68)80003-2

Vettin, J., and Todt, D. (2004). Laughter in conversation: features of occurrence and acoustic structure. *J. Nonverbal Behav*. 28, 93–115. doi: 10.1023/B:JONB.0000023654.73558.72

Vettin, J., and Todt, D. (2005). Human laughter, social play, and play vocalizations of non-human primates: an evolutionary approach. *Behaviour* 142, 217–240. doi: 10.1163/1568539053627640

Vick, S. J., Waller, B. M., Parr, L. A., Smith-Pasqualini, M. C., and Bard, K. A. (2007). A cross species comparison of facial morphology and movement in humans and chimpanzees using FACS. *J. Nonverbal Behav*. 31, 1–20. doi: 10.1007/s10919-006-0017-z

Waller, B. M., Caeiro, C. C., and Davila-Ross, M. (2015). Orangutans modify facial displays depending on recipient attention. *PeerJ*. 3:e827. doi: 10.7717/peerj.827

Waller, B. M., and Cherry, L. (2012). Facilitating play through communication: significance of teeth exposure in the gorilla play face. *Am. J. Primatol*. 74, 157–164. doi: 10.1002/ajp.21018

Waller, B. M., and Dunbar, R. I. (2005). Differential behavioural effects of silent bared teeth display and relaxed open mouth display in chimpanzees (*Pan troglodytes*). *Ethology* 111, 129–142. doi: 10.1111/j.1439-0310.2004.01045.x

Waller, B. M., Whitehouse, J., and Micheletta, J. (2017). Rethinking primate facial expression: a predictive framework. *Neurosci. Biobehav. R*. 82, 13–21. doi: 10.1016/j.neubiorev.2016.09.005

Weigel, R. M. (1979). The facial expressions of the brown capuchin monkey (*Cebus apella*). *Behaviour* 68, 250–276. doi: 10.1163/156853979X00331

Wild, B., Rodden, F. A., Grodd, W., and Ruch, W. (2003). Neural correlates of laughter and humour. *Brain* 126, 2121–2138. doi: 10.1093/brain/awg226

Wildman, D. E., Uddin, M., Liu, G., Grossman, L. I., and Goodman, M. (2002). Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus Homo. *Proc. Natl. Acad. Sci. U.S.A*. 100, 7181–7188. doi: 10.1073/pnas.1232172100

Winkworth, A. L., Davis, P. J., and Adams, R. D. (1995). Breathing patterns during spontaneous speech. *J. Speech Hear. Res*. 38, 124–144. doi: 10.1044/jshr.3801.124

Wood, A., and Niedenthal, P. (2018). Developing a social functional account of laughter. *Soc. Pers. Psychol. Compass* 12:e12383. doi: 10.1111/spc3.12383

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 23 July 2021
doi: 10.3389/fpsyg.2021.675172

Aping Language: Historical Perspectives on the Quest for Semantics, Syntax, and Other Rarefied Properties of Human Language in the Communication of Primates and Other Animals

Drew Rendall*

Department of Biology, University of New Brunswick, Fredericton, NB, Canada

In 1980, Robert Seyfarth, Dorothy Cheney and Peter Marler published a landmark paper in *Science* claiming language-like semantic communication in the alarm calls of vervet monkeys. This article and the career research program it spawned for its authors catalyzed countless other studies searching for semantics, and then also syntax and other rarefied properties of language, in the communication systems of non-human primates and other animals. It also helped bolster a parallel tradition of teaching symbolism and syntax in artificial language systems to great apes. Although the search for language rudiments in the communications of primates long predates the vervet alarm call story, it is difficult to overstate the impact of the vervet research, for it fueled field and laboratory research programs for several generations of primatologists and kept busy an equal number of philosophers, linguists, and cognitive scientists debating possible implications for the origins and evolution of language and other vaunted elements of the human condition. Now 40-years on, the original vervet alarm call findings have been revised and claims of semanticity recanted; while other evidence for semantics and syntax in the natural communications of non-humans is sparse and weak. Ultimately, we are forced to conclude that there are simply few substantive precedents in the natural communications of animals for the high-level informational and representational properties of language, nor its complex syntax. This conclusion does not mean primates cannot be taught some version of these elements of language in artificial language systems – in fact, they can. Nor does it mean there is no continuity between the natural communications of animals and humans that could inform the evolution of language – in fact, there is such continuity. It just does not lie in the specialized semantic and syntactic properties of language. In reviewing these matters, I consider why it is that primates do not evince high-level properties of language in their natural communications but why we so readily accepted that they did or should; and what lessons we might draw from that experience. In the process, I also consider why accounts of human-like characteristics in animals can be so irresistibly appealing.

Keywords: language, evolution, animal communication, primates, semantics

Edited by:
Marina Davila-Ross,
University of Portsmouth,
United Kingdom

Reviewed by:
John Lauderdale Locke,
City University of New York,
United States
D. Kimbrough Oller,
University of Memphis, United States

****Correspondence:***
Drew Rendall
d.rendall@unb.ca

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *02 March 2021*
Accepted: *21 June 2021*
Published: *23 July 2021*

Citation:
Rendall D (2021) Aping Language:
Historical Perspectives on the Quest
for Semantics, Syntax, and Other
Rarefied Properties of Human
Language in the Communication
of Primates and Other Animals.
Front. Psychol. 12:675172.
doi: 10.3389/fpsyg.2021.675172

THE VERVET ALARM CALL STORY AND ITS ENDURING LEGACY

Seyfarth et al. (1980a) published a landmark paper in *Science*, reporting what was interpreted to be evidence for language-like communication in vervet monkeys, a species of primate relatively distantly related to humans. The paper reported that vervet monkeys gave acoustically distinct alarm calls to different types of predator which prompted functionally distinct escape responses in listeners. It was argued that the calls were not simply emotively based and that contextual details were not needed in order for listeners to respond appropriately. Rather, the calls alone were sufficient to elicit the distinct escape responses. Hence, the calls appeared to function as symbolic labels for the predators, much like our human words for them, and were interpreted as the first evidence for semantic communication in a primate. Because the alarm vocalizations showed no iconic resemblance to the predators themselves, they were also claimed to exemplify the property of *arbitrariness* that the linguist Saussure had previously proposed to be a defining structural property of human words (de Saussure, 1971). Here then appeared to be evidence for language-like communication in a non-human primate with the potential also for some similar human-like cognitive abilities. The implications for the evolution of language and mind in humans – topics that had bedeviled scholars for ages – were tantalizing.

Indeed, the impact of the 1980 *Science* paper was profound. Although its findings were never replicated (until they couldn't be: Price, 2013; Price et al., 2015), the paper nevertheless became the textbook example of language-like communication in animals and catalyzed a successful career research program for its primary authors, Seyfarth and Cheney focused on other evidence of human-like behavior and cognitive abilities in primates. Much of that research program was summarized for a wider audience in two successful popular books entitled, *How Monkeys See the World: Inside the Mind of Another Species (1990a)*, and *Baboon Metaphysics (2008)*, the latter title a nod to Darwin who suggested that philosophers of mind at the time, like Locke and others, would get more traction on the problem of human psychology (metaphysics) by studying baboons (as Seyfarth and Cheney indeed did: Cheney and Seyfarth, 2008).

The vervet alarm call story also catalyzed countless other studies searching for rudiments of semantics, and then also syntax and other rarefied properties of language, in the communication systems of other primates and a variety of non-primate species besides. And it served also to bolster a parallel historical tradition attempting to teach symbolism and syntax in artificial language systems to great apes.

In fact, it is difficult to overstate the impact of the original vervet alarm call story, for it helped to fuel field and laboratory research programs for several generations of primatologists, right up to the present, and kept busy a significant number of philosophers, linguists, and cognitive scientists debating possible implications for the origins and evolution of language and other vaunted elements of the human condition (e.g., Dennett, 1983; Premack, 1985; Bickerton, 1992; Pinker, 1994; Deacon, 1998; Hauser et al., 2002; Pinker and Jackendoff, 2005; Fitch, 2010; Berwick and Chomsky, 2016).

Now 40-years on, the original claims for semanticity in vervet alarm calls have recently been recanted (Price, 2013; Price et al., 2015; Seyfarth and Cheney, 2017) and additional evidence for symbolism, syntax, or other high-level intentional and informational properties of human language in the natural communications of non-humans is thin (Wheeler and Fischer, 2012; Scott-Phillips, 2015; Fischer and Price, 2017).

In hindsight, these outcomes might have been anticipated given other standard features of communication in non-human primates, including: that they have relatively small repertoires of different calls and use them in a wide range of contexts with little context-specific usage suggestive of discrete messages; that most calls are graded variants on a few basic structural themes of coos, grunts, barks and screams; that there is little cortical control of vocal production which is instead largely limbically driven and closely tied to emotions; that there is a conspicuous absence of social-cognitive intentionality in communication or other aspects of their behavior; and that there is little evidence of productive vocal learning; all of which are hallmarks of human speech and language (reviewed in Owren and Rendall, 2001; Penn and Povinelli, 2007; Hammerschmidt and Fischer, 2008; Jürgens, 2008; Hage, 2010; Owren et al., 2010; Rendall and Owren, 2013; Fischer and Price, 2017; Nieder and Mooney, 2020; Fischer, 2021).

Ultimately, we are forced to conclude that, although there may be some superficial resemblances, there are simply few substantive precedents in primates, or other species, for the high-level intentional, informational and representational properties of language, nor its complex syntax.

This conclusion was, in fact, reached by Cheney and Seyfarth (1998, 2005) themselves some time ago (1998, 2005) with their conclusion that primate communication is fundamentally not intentional the way language is: "*non-human primates' inability to represent the mental states of others makes their communication fundamentally different from human language*" (Cheney and Seyfarth, 2005, p. 135). These conclusions have been echoed and extended by other prominent researchers in the field. For example, Marc Hauser, formerly a notable figure in this line of research, concluded that: "*Although 40 years of research have been invested in the capacity of animals to produce or comprehend externalized symbols, the relevant evidence that they do so is, at best, weak*" (Hauser, 2009, p. 194). In a subsequent review of the state of research on language evolution, Hauser et al. (2014) later concluded that, "*Animal communication systems have thus far failed to demonstrate anything remotely like our systems of phonology, semantics, and syntax.*" Michael Tomasello, in his book, *Origins of Human Communication*, concluded that: "*Primate vocal displays are basically no different from those of other mammals. Vocal displays are mostly unlearned, genetically fixed, emotionally urgent, involuntary, inflexible responses… How could such mechanical reflexes be a direct precursor to any of the complexities of human communication and language…?*" (Tomasello, 2008, p. 53).

Revealingly, these conclusions have done little to slow the pace of research in this area. To wit, at the close of 2020 (and according to Google Scholar), the original 1980 *Science* paper on vervet alarm calls had been cited almost 1,500 times, and a companion paper published the same year in the journal *Animal Behavior*, also under the banner of semantic communication (Seyfarth et al., 1980b), had been cited just over 1,000 times – with no sign that the rate of citation of either paper has slowed since most of the above-noted conclusions were reached. On the contrary, the rate of citation for both papers is actually higher post-2000 compared to before. The authors' two popular books are even more widely cited. In early 2021, *How Monkeys See the World* had been cited 3,795 times and *Baboon Metaphysics*, published only in 2008, has already been cited 974 times. When one considers that the thousands of researchers who have cited these various works have themselves likely also been cited by hundreds, possibly thousands, of other researchers – a calculus of spread now familiar to us all in the midst of the global COVID-19 pandemic – it is obvious just how widely impactful the vervet alarm call story has been. Clearly, its appeal was and remains strong, and its influence has spread broadly and deeply, such that it could be some time before news of its revision reaches the diverse literatures where it has taken hold.

Note that although the weight of evidence now does not support the conclusions of the original *Science* paper, nor most other language properties it spurred the study of in other species, this outcome does not in any way represent a critique of the intelligence, achievements or inherent worth of any of the species studied in the process. This is a very important point that I'll return to later. Likewise, the conclusion also does not mean that there is no constructive continuity between primate and human communication – in fact there is continuity and it includes: some similar elements of basic vocal anatomy and basic processes of vocal production and thus similarity in the resulting sounds produced; some similarity in the peripheral mechanisms of vocal perception; some flexibility in call production and usage; a role for feedback in shaping infants' vocal production development; and a role for facial and other gestures in complementing vocal communication. Together these areas of overlap may point to some basic common building blocks of communication that could be part of the scaffolding for human communication, ultimately including speech and language, even if they do not yet illuminate much about the emergence of higher-level properties of language, such as its semantics and syntax. There are many constructive reviews of this evidence with suggestions for where future research could productively focus (e.g., Rendall et al., 1998, 2004, 2005; Fitch, 2000, 2020; Davila Ross et al., 2008, 2010; Ghazanfar and Rendall, 2008; Takahashi et al., 2015; Griebel et al., 2016; Boë et al., 2017; Nielsen and Rendall, 2018; Pomberger et al., 2018; Ghazanfar et al., 2019; Oller et al., 2019; Dezecache et al., 2020; Locke, 2021 this issue). Hence, I will not dwell further on that evidence here to rehearse what is well covered elsewhere.

Instead, I will focus in this article on two other broad and important questions that are prompted by the enduring legacy of the vervet alarm call story but that have never before been asked or addressed: **First, why are core properties of language, in fact, not manifest in the natural communications of non-human primates? And second, why did we ever think they should be?**

WHY ARE SEMANTICS AND SYNTAX NOT MANIFEST IN PRIMATE COMMUNICATION?

The Role of Intentionality

On the one hand, the answer to this question is quite straightforward, because there is a natural organizational hierarchy inherent in the semantic and syntactic properties of language that is grounded in its underlying intentionality. Thus, in human language, canonical speech acts are predicated on implicit (and sometimes explicit) mental state attributions about our audience, namely that they have thoughts or beliefs that are *about* the world that we want to engage by communicating with them (Grice, 1957). For example, they might think X, and we'd like to affirm for them our own understanding of X, or change theirs. Or we may think that they do NOT have any such knowledge about X but should. And so we tell them about X. Either way, our capacity for thoughts that are *about* things and for attributing the same capacity to others – to viewing them, like ourselves, as mental agents with internal states of knowledge or belief that are also about things – is referred to formally as *intentionality* (sensu Brentano, 1874). More prosaically, intentionality represents the cognitive impetus *to inform* others (or affirm, change, influence, or otherwise engage, their mental states). Of course, this informing function of language is only part of how language works. Nevertheless, it is the essential foundation for the canonical semantic and syntactic properties of language which have been the focus of parallel comparative research on animals, where semantics represents the conceptual content of all that informing, and syntax represents the higher-order organizational rules that emerge with a need to organize more complex semantics.

For most (possibly all) non-human primates, formal intentionality – mental state attribution – appears to be lacking (Cheney and Seyfarth, 1998, 2005; Penn and Povinelli, 2007; Call and Tomasello, 2008). Chimpanzees appear to understand the goal or purpose of another individual's behavior in instrumental terms – at least in controlled settings interacting with human partners – but they do not appreciate the mental states that lie behind others' behavior and how that affects what they will do, nor do they act deliberately to alter those mental states (Call and Tomasello, 2008). Studies of other primates have confirmed a similar lack of appreciation for others' perspectives or mental states, including specifically in communication. For example, Cheney and Seyfarth (1990b) explained early on that, while a vervet monkey will produce alarm calls when it perceives itself to be in danger, that same individual fails to call on other occasions when it is not in danger itself even though other companions, including kin, are at risk. Cheney and Seyfarth (1990a) subsequently replicated this finding with Japanese and rhesus macaques in a series of controlled experiments in captivity designed specifically to systematically test intentionality in communication. In that work, they found that mothers likewise failed to warn their infants of an imminent danger if the situation did not also represent a threat to the mother herself, and they concluded that this was because mothers failed to appreciate the perspective and knowledge of their infants when it differed from their own. Similarly, in field experiments on baboons, Cheney et al. (1996) showed that adult females routinely produce very loud bark vocalizations – informally termed "lost calls" – when they get separated from companions during daily travels in an effort to re-establish contact with the group. However, those same females do not respond vocally to the "lost calls" of other group members who become separated and are trying to relocate the group if they themselves are now safely with the rest of the group at the time. This failure to vocalize to inform others of the group's location held true even when the separated and calling group member was a female baboon's own young infant (Rendall et al., 2000).

Taken together, the corpus of work on primate intentionality consistently shows that the animals often attend carefully to the behavior of others, but fail to appreciate what lies behind it. They fail to appreciate others as mental agents with perspectives and resulting states of belief or knowledge that can differ from their own and that in turn affect their behavior. Hence, they fail also then to appreciate how their own behavior, including their own vocalizations, might serve to inform others – i.e., to change their states of belief or knowledge, and thus also their behavior. Hence, they vocalize when they themselves encounter a predator, find food, or become lost but this calling is entirely self-centered, reflecting their own current situation and needs. It does not reflect the informational needs of receivers. In contrast, human language is fundamentally "other-centered" in being routinely tailored to the perceived informational needs of listeners (Owren and Rendall, 2001).

Ultimately then, where formal intentionality in communication is lacking, there is simply no need or capacity for semantics or syntax. In other words, lacking the fundamental underlying cognitive impetus to inform others based on an appreciation of their states of knowledge or belief about the world, there is no functional need in non-human primates for any conceptual informing content to begin with, and therefore also no need for a higher-order syntactic system to organize more complex messages.

It is worth noting here that, in view of the negative findings on primate intentionality, an alternative conception of primate semantics arose in the 1990's that was referred to as "functional reference" and promoted continued work on the subject (Marler et al., 1992; Macedonia and Evans, 1993; Evans, 1997). The proposal was that, although primates were evidently not vocalizing intentionally to inform others about things in the world, as is the case for routine human language use, their vocalizations might nevertheless function "as if" they were. For example, while a vervet monkey producing an alarm call might do so with respect to its own circumstances and its own associated concern, fear or distress on encountering a predator, as is now accepted to be the case (Price et al., 2015; Seyfarth and Cheney, 2017), other group members hearing the call might still respond appropriately "as if" the call had conveyed semantic information about the predator type. In the original formulation of the framework, the threshold for assigning functional reference was that the vocalization elicited appropriate responses from listeners in the absence of supporting contextual information (Macedonia and Evans, 1993). By this criterion, vervet alarm calls would now not qualify even for this looser characterization of reference, given that listeners are now acknowledged to require additional contextual details to respond appropriately (Price, 2013; Price et al., 2015; Seyfarth and Cheney, 2017).

Latterly, however, this context-free criterion seems to have been relaxed further to allow signals, such as vervet alarm calls, to be labeled functionally referential if listeners can respond appropriately using additional information available from the immediate contextual details as well as what they have learned from past experience about the kinds of events that are associated with specific vocalizations from companions (Seyfarth and Cheney, 2017). In this way, what could be largely affective or motivationally driven vocalizations in signalers might nevertheless be interpreted to retain an element of external reference by virtue of additional inferences listeners make based on other information they glean from the current situation or past experience. On the surface of it, this is a perfectly sensible parsing of how the monkeys actually behave and respond. After all, primates are large-brained animals with significant inferential capacities. So, almost certainly, they routinely respond to vocalizations from companions based on a variety of inferences they make using a combination of immediate contextual details, their familiarity with group members and their individual behavioral proclivities, and their familiarity also with the circumstances that typically elicit different kinds of signals from them (cf. Smith, 1977; Owings and Morton, 1998).

At the same time, however, if vervet alarm calls function in this manner then they are effectively no different than any of the other vocalizations in their repertoire, or in the vocal repertoire of any other species, such as the common grunts, barks, squeals and screams that mediate quotidian activities. All of these calls too are largely affectively driven and reflect the current situation and needs of the signaler but could nevertheless similarly allow listeners to draw additional inferences about likely eliciting circumstances based on additional contextual information and familiarity with each other and their general proclivities.

Indeed, such a parsing of the function of vervet alarm calls simply aligns them with a wide range of other signaling phenomena not typically considered language-like at all, including, for example, human infant crying (and "crying" in other species). While the crying of human infants is definitively emotionally driven, parents can nevertheless often infer some general things about the eliciting circumstances, such as whether the cries reflect being overtired, or hungry, or in pain or general distress based on familiarity with their own infant, its crying patterns, and other contextual details including the time since last feeding, or the infant's recent sleep history (reviewed in Zeifman, 2001; Soltis, 2004). As a result, there does not appear to be any explanatory value added by attaching the label "reference" to vervet alarm calls, or any of these other signals, in order to promote comparison to language if the label could be applied equally well to all of these other common signals that are so clearly not language-like at all.

These points were well appreciated by earlier key researchers (see Premack, 1972; Marler, 1977; Owings, 1994), who therefore recognized an important distinction between such signaler-centered, affectively based vocalizations (in animals or humans) and the truly referential or symbolic quality of human language. The key distinction, as noted earlier, is that linguistic reference hinges on the intentionality of language users. That is what moves language beyond being exclusively sender-centered to being also receiver-centered, because language acts are routinely conditioned not only by the immediate circumstances of the speaker but also, and specifically, the informational needs of listeners. It is also what confers the representational power of language, allowing it to move beyond contextually bound signals that can be interpreted only with additional details of the immediate circumstances, to the context-free and virtually unbounded representational universe of human language instantiated by words that have common representational value for signaler and receiver alike and are understood by both parties to have such representational value.

In short, without intentionality, communicative acts have no meaning in the formal linguistic sense, and cannot be scaffolded into more complex semantic constructs that create pressure for organizational systems (grammars) to organize them. So it is precisely the psychological characteristic of formal intentionality explained earlier that kicks off the complexity of linguistic reference and that an evolutionary account of language therefore needs to account for. It is, therefore, illusory to search among primates, or other animals, for vocalizations that are referential only indirectly in some functional and not intentional sense, maintaining that this will somehow provide any illumination on the evolution of linguistic reference (reviewed in Rendall et al., 2009; Owren et al., 2010; Rendall and Owren, 2013).

So, while understandable enough as a conceptual retreat regarding primate semantics in response to mounting evidence for a lack of language-like intentionality in their communication, the functional reference gambit actually muddies comparisons between language and primate communication and obfuscates more than it illuminates meaningful points of similarity and difference between them that could ultimately clarify our understanding of the course of language evolution. In a recent comprehensive review of the concept of functional reference, Wheeler and Fischer (2012) drew much the same conclusion, allowing that the functional reference framework was "a promising paradigm whose time has passed."

In summary then, the straightforward answer to why non-human primates lack semantics and syntax is that they lack the functional prerequisites to each: they lack syntax because they lack the prerequisite complex semantics, and they lack semantics because they lack the prerequisite intentionality. There is a functional hierarchy to these properties of language, where, for non-human primates, the ground floor is missing.

It is very important to appreciate the contingent functional nature of these language properties, and why they are then absent in non-human primates. At the same time, however, this understanding seems just to push the matter back one step. **Why would these canonical features of language not also be important to non-human primate communication?**

The Natural Environment of Primate Communication

That's a different and important question – the "other hand" of the issue as noted above. Currently, it's impossible to say definitively why primates do not manifest semantics or syntax in their natural communications because the question has never been explicitly posed and studied in that way. However, traction on the question is likely to come from refocusing on the natural history, environment and behavior of the animals themselves – rather than seeing them as stand-ins for human ancestors – and assessing how their communication in fact serves the needs of their world, rather than ours. In this, there is a wealth of relevant literature to draw on from the significant bodies of research on primate behavior and ecology generally, and on general aspects of their communication unrelated to the question of language. And, while Primates are a large and diverse Order, there are some common elements that stand-out in these literatures and provide fertile ground for addressing the question (Smuts et al., 1988; Mitani et al., 2012).

Very briefly then, many primate species are highly social, even those that might typically be labeled as "solitary" because they do not live in permanent groups. There are important exceptions, of course, but to the extent broad generalization is possible, and despite many variations in other details of their behavior and ecology, primates are generally held to occupy a distinctly social niche. Many species live in stable groups, or looser communities, comprised of individuals of varying age, sex, social rank, and degrees of relatedness. And most species are also relatively long-lived. Hence, there is protracted opportunity for development of a complex web of differentiated social relationships among group members according to these various social distinctions. Indeed, the conclusion from many decades of research is that these different social relationships powerfully affect all manner of daily activities and have resulted in a highly developed "social acumen" (Jolly, 1966; Humphrey, 1976; Dunbar, 1998).

Many daily activities are also mediated by vocalizations, including soft coos and grunts that mediate relatively relaxed, affiliative social contexts; loud barks and screams that mediate aggressive conflict; and excited squeals and shrieks in the context of food discovery or predators. Many of these contexts have a particular immediacy to them and are strongly valanced, either positively or negatively, with variation in the intensity thereof. And, while innately given, the vocalizations produced in these contexts also reflect that variable valencing of events through marked grading in the amplitude, tempo and spectral structure of the calls in any particular circumstance. Vocalizations produced in some of these contexts also manifest structural differences attributable to a variety of indexical dimensions, namely differences in the age, sex, size and often also individual identity of the caller. Hence, the social attunement of the animals and their rich and differentiated social histories provides broad scope for inference and interpretation of such affectively laden signals – as noted earlier – according to the identity of the caller and their age, sex, rank, and kinship relative to listeners; further conditioned by the recent and longer-term nature of their relationships to one another; and by myriad elements of the immediate behavioral and environmental context associated with calling. Taken together, this mix of cues from available contextual details, the social identity of signalers, and the dynamic social history of participants describes a pretty rich platform for flexible and functional communication in support of a host of quotidian social and behavioral routines. But not one where there is an obvious selective need for anything like the semantic or syntactic properties of language.

This is a necessarily truncated parsing of primate communication for present purposes and is not intended to decide the question of why primates do not evince semantics or syntax but rather only to open it. Nevertheless, it is a parsing that aligns with many earlier proposals (Smith, 1977; Krebs and Dawkins, 1984; Premack, 1985; Leger, 1993; Owings, 1994; Owings and Morton, 1998; Owren and Rendall, 2001; Wheeler and Fischer, 2012) and also with very recent and comprehensive reviews of the subject (Fischer and Price, 2017; Fischer, 2021). Notably, it is a conclusion endorsed recently also by Seyfarth and Cheney (2017). In revisiting the original vervet alarm call work in the light of their own recent revised findings, they conclude together with colleagues that: "*We suggest that both cognitive appraisal of the situation and internal state contribute to the variation in call usage and structure. While the semantic properties of vervet alarm calls bear little resemblance to human words, the existing acoustic variation, possibly together with additional contextual information, allows listeners to select appropriate responses*" (Price et al., 2015, page 1). And in a more recent broader review of the topic, Seyfarth and Cheney (2017) now emphasize the constrained nature of call production in primates that is largely innately given and tied importantly to affective motivations, making contextual details important to listeners in interpreting and responding to the vocalizations of others in any given situation. These recent acknowledgments represent a pretty significant reversal of perspective from having originally and explicitly discounted the importance of affect and contextual details in vervet alarm calls in favor of the semantic quality of the calls alone based on categorically distinct acoustic structures and specific contexts of usage (Seyfarth et al., 1980a). Notably now, however, they reinterpret the above-noted characteristics of primate communication under the banner of linguistic *pragmatics* rather than *semantics*. So, with this revised perspective, they continue to attempt concrete connections to language even while acknowledging that the previous focus on semantics was misplaced. **Which prompts again the question, why we ever expected primate communication to evince semantic or syntactic properties similar to human language to begin with?**

WHY DID WE EVER THINK NON-HUMAN PRIMATE COMMUNICATION SHOULD SHOW LANGUAGE-LIKE SEMANTICS OR SYNTAX TO BEGIN WITH?

On the one hand, it is a perfectly sensible intuition that primates might manifest some of the rudiments of human communication, possibly including language. They are, after all, our closest living relatives. Hence, it is entirely reasonable to study primate communication for elements of it that might inform our understanding of the origins and evolution of language. On the other hand, it might feel a backward, or at least a bit strained, to search first specifically for evidence of the high-level intentional and representational properties of language, such as its semantics and syntax, if these are at all likely to be relatively recent, derived and specialized properties of communication in humans. In which case, it might feel strained to expect meaningful precursors of such specialized language properties in species so far removed from modern humans. It may even feel logically backward to effectively project such derived properties of modern language backward in time to the communication systems of living species as stand-ins for human ancestors assuming these properties of language must also be present and functional for them as well, even if in more rudimentary form (Premack, 1985; Pinker, 1994; Rendall et al., 2009).

Whatever one's stance, it should be noted that scholarly and popular interest in the possible evolutionary precursors to language in primate communication long predates the vervet alarm call story. Up until that time, however, research on the subject was relatively spotty and fragmented. **So, the pertinent issue for present purposes is really why the vervet alarm story was so especially impactful and how it served to consolidate a much more focused and enduring research agenda?**

This is an important and potentially multi-faceted question. Part of the answer may lie in effects attributable to the historical, intellectual climate of the time – and a paradigm shift in that – as well as to the specific intellectual commitments of key players involved in that paradigm shift. Another part may involve broader influences on the wider audience that affected their reception to, and uptake of, the original findings.

WHAT MOTIVATED THE FOCUSED SEARCH FOR SEMANTICS AND SYNTAX THAT ACCOMPANIED THE VERVET ALARM CALL STORY?

The Rise of Cognitivism

One can trace the roots of this focused search to a couple of parallel scientific developments of the mid-20th century. The first of these was the rise of cognitivism in Psychology which involved shedding the shackles of Behaviorism. Importantly, this paradigm shift included a specific focus on language and also included a very public showdown between Chomsky and Skinner concerning the extent to which behaviorism or cognitivism represented the better approach for understanding human language (and behavior more broadly). Chomsky was the decided winner and, while certainly not alone in this, was a central figure in helping to usher in a cognitive revolution that took hold in Psychology and ultimately transformed many disciplines, and even invented some entirely new ones (e.g., Cognitive Science) all with a fresh focus on human mental experience. Many researchers in animal behavior were also quick to embrace cognitivism. After decades of the strictures of Behaviorism in the study of animals as well, they too were poised to think again about animal mental life. This focus ultimately led to a reorientation and rebranding of a whole branch of animal behavior research under the banner of Cognitive Ethology.

A second important and complementary development involved research at about the same time specifically on animal communication. A key early figure in this development was Donald Griffin at Rockefeller University who, together with others, was responsible for solving the mystery of how bats navigate in the dark – namely through production of a continuous stream of high-frequency (ultrasonic) clicks and detection of their reflected echoes off objects in the environment, a process dubbed echolocation. This was an exciting finding that helped to illuminate (for us) the dark world of the bat. For Griffin, it also captured how communication, among all behaviors, could be a privileged source of insight into animal mental experience – a "window into their minds" as he put it (Griffin, 1995). He promulgated this notion for a wider audience in a popular book entitled *The Question of Animal Awareness (1976)*, and another titled, *Animal Thinking (1985)*, in which he wholly re-popularized the formerly taboo subjects of animal mental experience and animal consciousness under the new banner of Cognitive Ethology. In the first of these books, he also explicitly forged the connection to language in a way that would help frame the subsequent conceptual agenda for much comparative research on primates, writing that:

> "*In so far as linguists and philosophers have been correct in linking human thinking so closely to language, the communication behavior of other species is bound to suggest conscious thought to roughly the extent it shares essential features with human speech*" (Griffin, 1976: p. 39).

The Rockefeller Effect

These various threads were tied together by Peter Marler, who was by then a colleague of Griffin's at Rockefeller, having relocated earlier from Berkeley. Marler's career research program – prior to this and following – was focused primarily on birdsong, though with an abiding interest in comparisons to language. Significantly at the time, though, he had recently supervised a stand-out graduate student named Tom Struhsaker at Berkeley. Struhsaker had conducted a comprehensive field study of vervet monkeys and published a monograph on his research (Struhsaker, 1967). It focused primarily on the natural history of the monkeys, general dimensions of their behavior and ecology, but it also included a section on communication. Importantly, that section contained preliminary descriptions of a small repertoire of alarm calls produced in reaction to different kinds of predators. Struhsaker did not pursue the matter in detail. However, for Marler, the combination of Griffin's local influence at Rockefeller and the specific research challenge Griffin had laid down connecting animal language to conscious thought, assuredly left Struhsaker's brief descriptions of the vervet alarm calls pregnant with possibility.

At Rockefeller then, Marler recruited as postdocs Robert Seyfarth and Dorothy Cheney fresh from completing PhDs on the social behavior of baboons in South Africa, and he dispatched them to Kenya to followup Struhsaker's preliminary descriptions of vervet alarms. Marler also recruited two other teams to conduct similar studies on other primate species, one of them another husband and wife couple, Harold and Sally Gouzoules, and the other, Steven Green. Notably, all three teams returned reports of categorically distinct vocalizations interpreted to manifest parallels to the semantic properties of language, the first ever such reports from the natural communications of primates.

Seyfarth and Cheney returned the now familiar vervet alarm call story (Seyfarth et al., 1980a). Further details of their fieldwork and how they came to their conclusions is treated thoroughly in an engaging book by the historian of science, Gregory Radick (Radick, 2007). The Gouzoules' studied loud scream vocalizations given by many primate species when physically attacked by social companions. They returned evidence of distinct variants of scream in rhesus monkeys that were proposed to convey a host of representational information to listeners about the severity of the aggression involved, the social rank of the attacker, and the degree of kinship between attacker and victim (Gouzoules et al., 1984). And Steven Green studied "coo" vocalizations produced by Japanese monkeys in a variety of social contexts and provided a typology of different kinds of coo which were reported to be perceived and interpreted categorically much like the sounds of human speech (Green, 1975; Zoloth and Green, 1979). Neither of the latter two studies garnered quite the same attention as the vervet alarm call story, nor were they ever replicated, but collectively they helped catalyze an enduring tradition of language parallels research in primates and beyond.

Much of this enduring tradition of language parallels research, particularly that focused on primates, continued to be closely connected to Marler's trainees, such as Seyfarth and Cheney, and then their trainees, and the trainees of those trainees in turn, each focused on additional evidence of semantics, and then also syntax, in the vocalizations of primates (e.g., Gouzoules et al., 1984; Evans et al., 1993; Macedonia and Evans, 1993; Hauser, 1998; Zuberbühler et al., 1999; Manser et al., 2002; Slocombe and Zuberbühler, 2007; Arnold and Zuberbühler, 2008; Crockford et al., 2015; Zuberbühler, 2019). So, there is a very close connection of this research tradition to a family of researchers emanating from Marler and Rockefeller University.

At the same time, the research tradition forged by this group has also spread much more widely and ultimately taken in a wide range of other researchers and a broad array of animal species from chickens and chickadees, to meerkats and marmosets, to prairie dogs and squirrels (e.g., Evans et al., 1993; Greene and Meagher, 1998; Manser et al., 2002; Templeton et al., 2005;

Kitzmann and Caine, 2009; Slobodchikoff et al., 2009). It is important to ask then what accounts for this wider appeal? How did this line of research captivate such a broad audience?

WHY WAS THE WIDER AUDIENCE SO RECEPTIVE TO THE VERVET ALARM CALL STORY?

The citation statistics for the vervet alarm call story noted earlier are striking, particularly when one appreciates that those citations represent only the record of formal published works explicitly influenced by the original study. The more informal influence of the original story on scholarly thought and research programs must be far more extensive.

This broad influence is the more remarkable when one considers that the original work was only a single, unreplicated study. And more remarkable still when one appreciates that the original evidence was also quite mixed. As reported in the original papers, the alarm calls were not, in fact, given exclusively to predators but were used also in other contexts characterized by high arousal such as within group aggression; there were acoustic gradations among the various calls; and the monkeys' responses to hearing different alarm variants were also mixed. Indeed, only some of the main effects tested experimentally in the original *Science* paper were statistically significant and then only very few using a conservative alpha level of *0.01*. The other effects interpreted as significant were subject to relaxed alpha levels between *0.05* and *0.10*. So, the production and response patterns for the different alarm calls were never either definitive nor exclusive.

This mixed pattern of call usage in both alarm and non-alarm contexts, and the overlap in call structures between the proposed alarm variants, has been quantified and confirmed much more thoroughly in a recent study deliberately revisiting the original vervet study (Price, 2013; Price et al., 2015) leading to the revised characterization of the alarm calls quoted earlier:

> "*We suggest that both cognitive appraisal of the situation and internal state contribute to the variation in call usage and structure. While the semantic properties of vervet alarm calls bear little resemblance to human words, the existing acoustic variation, possibly together with additional contextual information, allows listeners to select appropriate responses" (*Price et al., 2015, page 1).

However, this variability and interpretation was not emphasized nor widely appreciated at the time of the original study, nor since. Somehow the subtleties got lost in the re-telling of the work which came to emphasize the discrete, categorical nature of the calls to pick out different types of predator and thus their apparent language-like semantic properties. In fact, given that citations of the authors' two popular books exceed those of the original primary publications by a factor of 3:1, it is possible that many scholars who have cited the vervet work are not actually familiar with the original paper and its findings.

This is all the more remarkable given that the implicit, if not explicit, scholarly code in science is that – *extraordinary claims require extraordinary evidence*. This well-known prescription is originally attributed to the French polymath, Laplace, and was popularized in contemporary times by the cosmologist, Carl Sagan (Gillispie, 2000). It is a scholarly code well-known to researchers in evolutionary-oriented disciplines from the writings of George C. Williams, a luminary in evolutionary biology in the 20th century. Williams invoked the prescription in reference to claims of adaptation (e.g., Williams, 1966). He regarded adaptation as a weighty construct and indeed one that had to be if it were to have any real value in evolutionary theory (cf. S. J. Gould). Hence, he argued that claims of adaptation cannot be made lightly, nor accepted uncritically by others. Instead, claims of adaptation bear a heavy burden of proof. Researchers advancing claims of adaptation and readers evaluating them must be equally circumspect and dually committed to high standards of evidence.

Although Williams helped to codify this notion in evolutionary biology in the mid-20th century, it was appreciated by evolutionists well before that. Darwin himself might be the paradigm example. He had to be literally cajoled to publish his theory of evolution by natural selection and he took the better part of 30-years to muse on it and to amass the requisite evidence before releasing it in print. Darwin knew his "dangerous idea" (cf. Dennett, 1995) would seem an extraordinary claim; hence, it needed extraordinary evidence, and he took 30-years to carefully and comprehensively accumulate it. Darwin's diligence delivered significant dividends. Most of his core insights, some very far-reaching, have stood the test of time.

There feels like a lesson here for contemporary science, now conducted at break-neck pace by comparison and increasingly handmaiden to a host of additional factors quite peripheral to the science itself, including grant funding and prestigious awards, impact factors, citation statistics, media coverage, popular attention, and ultimately career advancement.

In another example from the vervet research, Seyfarth and Cheney reported an additional landmark finding in 1984 shortly after the alarm call study was published, this time in the journal *Nature* (Seyfarth and Cheney, 1984). The *Nature* paper reported evidence of tit-for-tat reciprocal altruism in the vervets, a form of cooperation virtually undocumented in animals at that time but thought to be central to the complexity of human social behavior and cooperation. The vervet reciprocity study was based on a sample of nine subjects, which is the minimum sample size required to achieve a significant effect by the non-parametric Wilcoxon test when one of the subjects responds counter to the hypothesis, as was the case. Another extraordinary claim based on thin evidence, now cited more than 600 times. This additional example might appear to be focusing on the vervet research, in particular, but that is not the point. There are likely myriad other similar examples in this and other fields. No doubt any reader of this can point to one or more similar studies in their own particular area of research. And *that* is the point, that some findings have a powerful appeal, absent the usual standard of evidence expected. So what accounts for that? **Why, despite the cautionary prescription about standards of evidence echoed across time by leaders like Laplace, Williams and Sagan, are we so credulous of certain findings? In particular, why are we so credulous of reports of human-like characteristics in animals,**

such as reports of semantic communication or reciprocal altruism?

Our Anthropomorphic Instinct

Possibly because they appeal to our anthropomorphic instinct, our habit of attributing human-like qualities to animals and even to non-material entities (e.g., gods, spirits, the weather, Mother nature, etc.). Non-human primates certainly look like us in many ways, and they can also move and act a lot like us, as well. So, it's natural to assume that they communicate and think like us, too. Of course, the latter assumption is fraught, particularly when continuity of mental experience serves then as both the *a priori* assumption guiding scientific enquiry as well its conclusion. Nevertheless, the prospect that vervet monkeys, and other species besides, might be a lot more like us than previously known may have been, and continue to be, irresistibly appealing.

Anthropomorphism is a long-standing and possibly universal human practice. Popular and scholarly stance on it has varied. In the middle ages, it was apparently common to put barn-yard animals on trial for bad conduct based on attribution to them of a moral sense (i.e., they should know better). Cartesian dualism subsequently swung the pendulum, proposing that humans were uniquely endowed with the ability to reason and reflect (to cogitate), while animals were driven by emotion and instinct. Animals were likened to mechanical automata, a stance that comfortably distanced us from their instinctual and seemingly brutish habits. And that was part of the cause for distress to Darwin in publishing his major treatise on natural selection: because his theory of evolutionary descent, that would have explicitly connected humans to a shared ancestry with apes and other primates, would surely offend Victorian sensibilities. To wit, the quote often attributed to the wife of the Bishop of Worcester speaking to her husband on hearing Darwin's idea: "My dear, descended from the apes? Let us hope it is not true. But if it is, let us pray, that it will not become generally known" (Leakey and Lewin, 1977, p. 21).

However, as Darwin's views did become more widely known and the thinking around evolution and biological continuity became more broadly established in the late 1800's, the pendulum swung again, back to attributing to animals more sophisticated, human-like forms of reasoning and consciousness, as epitomized in the work of George Romanes. Romanes was a friend and champion of Darwin and enthusiastically extended the biological continuity inherent in Darwinian descent to include continuity of psychological experience as well, in his founding of the new discipline of Comparative Psychology. Romanes was at pains to legitimize study of animal psychology, as a *bona fide* science, which he argued had too long been the purview only of amateurs. In his major treatise (Romanes, 1882: *Animal Intelligence*), he covered a wide range of mental and emotional phenomena across all major animal groups, from invertebrates to primates, inferring mental experience in animals using a combination of personal introspection and analogy. Specifically, he argued that, where the behavior or activity of animals was similar to that of humans, we can infer that the underlying mental operations are also equivalent and are specifically those revealed to us by our own introspection. He was quite explicit in championing this anthropomorphic method. Some of the work involved empirical study, but a lot of it was quite speculative and based on anecdote. And the excesses of this period in comparative study were part of what motivated Lloyd Morgan's eponymous canon, appealing to the principle of parsimony in application to psychology (i.e., never attribute to the action of a higher mental faculty behavior that can be adequately accommodated by a lower one), and motivated the broader subsequent behavioristic paradigm in comparative psychology in the early 20th Century.

With Thorndike, Watson, Skinner, and others, Behaviorism swung the pendulum once again to a focus explicitly and only on what was concretely observable and measurable – namely behavior – eschewing all reference to internal mental states. The backlash of Behaviorism marked significant progress in empirical methods and techniques for the study of behavior but was quite stultifying in its proscription of all things mental. And, so, by the mid-20th century, in another swing of the pendulum, there was an almost palpable release as the Cognitive Revolution prompted a spirited revival of interest in, and research on, human and animal mental life.

This cognitive revival was closely connected to a growing appreciation of Darwinian evolutionary principles and much greater acknowledgment, both popularly and in scholarly circles, of the connectedness of all living things. The idea of behavioral and cognitive continuity was no longer so threatening as it was in Victorian England when the notion of continuity created considerable dissonance by connecting us to what was assumed to be a brutish animal past. In fact, with attribution of much more human-like characteristics to primates and other animals, Darwinian evolutionary continuity may now be much more flattering in thereby also bequeathing us a more auspicious ancestry. Indeed, modern findings in genetics that reveal clear traces of Neanderthal ancestry in many contemporary human populations (Sankararaman et al., 2014), and evidence for Neanderthals' more sophisticated material culture (Hardy et al., 2020), have fundamentally changed and improved the popular image of these hominin representatives that were formerly held in very low regard.

Debate about the power versus pitfalls of anthropomorphism is longstanding (Kennedy, 1992; Eddy et al., 1993; Budiansky, 1998; Sober, 1998; de Waal, 1999, 2018; Povinelli et al., 2000; Wynne, 2004; Klopfer, 2005; Barrett et al., 2007; Rendall et al., 2007; Urquiza-Haas and Kotrschal, 2015). Those who support it see tremendous heuristic value and propose that it might even represent an adaptive form of human reasoning – an adaptive human instinct (Barrett, 2005; Urquiza-Haas and Kotrschal, 2015). Attributing human-like traits to fellow humans is obviously wholly natural. It might also be truly adaptive if, as seems plausible, there are fitness advantages to projecting one's own internal experience onto other humans as a way of "reading their minds" to better understand, anticipate, and influence their motivations and behaviors. Generalizing that strategy of "human projection and mind reading" to other species is also certainly habitual, at least in contemporary western society (cf. Gray, 2020). We routinely talk to our pets and think we understand them, and that they in turn understand us. We also talk to gods and other spirits in the hope that they are listening and will humor us. We see faces in the clouds, and we revel in cartoon animals that are made to talk, act and dress like humans. Indeed, there are whole genres of children's books, movies and television programs based on deliberately humanized animal characters (e.g., Winnie-the-Pooh; Yogi Bear; Smokey the Bear; Big Bird; Barney; Bagheera and Baloo, Tom and Jerry, O'Malley Cat, Wile E. Coyote, and any of a hundred other Disney and Hanna-Barbera characters). Hence, children in many western cultures, at the least, are enculturated into anthropomorphism, if the habit is not already instinctual. And so whether as adaptive human instinct, or as enculturated habit (or both), attributing human-like characteristics to animals, particularly those most similar to us in other respects such as primates, is common and possibly quite difficult to resist. It is also probably low cost, even if wrong, at least in most of these circumstances ... except possibly in scholarly endeavor.

Hence, those who caution against anthropomorphism warn that it is risky to assume commonality of mental experience with other species, the moreso as the phylogenetic distance from humans increases, even where there might be obvious similarity otherwise in external appearance and behavior (Eddy et al., 1993; Povinelli et al., 2000; Wynne, 2004; Barrett et al., 2007). A simple contemporary analogy makes the point: while the Tesla and the gas-powered Jaguar may look a lot alike, have a lot of the same peripheral hardware, and basically do the same functional things (both provide a fast, comfy ride), when you look under the hood, the way they get things done is fundamentally different (a combustion engine versus a really big battery . . . which isn't even located under the hood). Critics of anthropomorphism, at least as a scientific methodology, see the same problem viz a viz the internal mental engines "under the hood" that drive our behavior versus that of other species: they might be similar or they could be quite different.

Indeed, extending psychological continuity across broad taxonomic distances glosses a fundamental element of Darwin's evolutionary insight, namely that the so-called "tree of life" is truly a tree and not a ladder. The latter, ladder-like view of evolution was formerly quite popular, captured in the classic *Scala Naturae,* which envisioned evolution as a linear and progressive process culminating in humans (just short of God). However, that view was rightly abandoned with the Darwinian revolution that emphasized the diversifying effects of the evolutionary process and thus the diversified products it yields. Hence, we now understand species not as rungs on a ladder leading to humans, but rather as the tips of branches of a vast evolutionary tree, where each branch, including ours, describes a distinct path with potentially very distinct evolutionary challenges and solutions. So, it is dangerous to gloss that deep and diversifying evolutionary history and to treat contemporary species – the tips of many different branches – as stand-ins for human ancestors and so as scaled-down versions of ourselves (Hodos and Campbell, 1969, 1990).

This point is well understood by most evolutionary researchers, particularly those who study structure and form (e.g., anatomy) which leaves a fossil record that is tangible and concretely measurable. Here it is possible to trace the distinct evolutionary paths taken by different species through the tangible, measurable evidence of their ancestors. Admittedly, it's more difficult for researchers interested in mental experience, because mental activity does not fossilize so directly and it is also not so easily accessed or measured even in living descendants (at least not yet). Everything has to be inferred. So, what choice does one have other than to use extant species for comparison to humans? And then how does one go about studying their mental experience? On what do you base your hypotheses of how and what they think? Where is your objective, unbiased point of entry into their mental lives? It's hard enough to get inside the head of other human beings to truly appreciate their perspectives, thoughts, and feelings; it's much harder still to get inside the head of another species when their "Umwelt" and resulting "Innenwelt" (von Uexkuüll, 1909) may be so different. Therefore, making the anthropomorphic gambit – projecting our own internal experience onto that of other species – is risky. But what's the alternative? And so anthropomorphism, a natural human instinct or enculturated habit, practiced routinely and informally in everyday life, elides into a common research strategy in science as well it seems. This elision may often be unconscious or possibly tacit, however, because most scientists would probably deny that they ascribe to anthropomorphism formally and may not even be aware when they are, nevertheless, practicing it.

So, the appeal of the vervet alarm call story, and the broader research program it galvanized, may in part reflect our increasing comfort now with, and indeed the general appeal of, explanations of animals that emphasize their continuity with humans, whether or not we are even aware of our sympathy and appetite for this.

Conservation and Animal Rights

It's also possible that some of the appeal of the vervet monkey story, and other work in the genre, reflects the growing movement to engender greater understanding and compassion for other species, an outgrowth of the ever-increasing sense of eco-awareness and conservation that began in the mid-1900's and has grown steadily since (if too slowly, still), based on a broader understanding of, and commitment to, the inherent connectedness of global ecosystems and their inhabitants.

This sentiment, in more explicit and exaggerated form, finds expression in some of the research conducted on great apes that seeks parallels to language. One of the most celebrated of these concerns the long-running research program by Sue Savage-Rumbaugh with Kanzi, a bonobo (which is a species of chimpanzee, formerly but no longer referred to as the *pygmy* chimpanzee). Kanzi was reared from birth in a language research environment and was regularly immersed in human routines such as cooking meals and camping out in the forest. Ultimately, he learned a large number of artificial symbols. The work also garnered a lot of public attention and press, some good and some bad. It was popularized for a wide audience in a book co-written with the science writer, Roger Lewin, titled, *Kanzi: The Ape at the Brink of the Human Mind* (Savage-Rumbaugh and Lewin, 1994). Here and elsewhere, his tutor and advocate, Sue Savage-Rumbaugh, proposed that Kanzi be extended human rights. She even took the step of adding Kanzi as a co-author on research publications, apparently as *ipso facto* proof of his humanity (Savage-Rumbaugh et al., 2007).

Here, the effort to persuade others to extend human rights to Kanzi effectively turns on his mastery of language skills, the logic being that any creature possessed of sophisticated language abilities must be sentient in a very human-like way and thus must be accorded some of the same rights and privileges (an extension of Don Griffin's argument connecting language-like communication in animals with conscious thought). Put simply: If language, then human. To be sure, the objective here is laudable – but it might not be objective. Certainly, it must put added pressure on the research program, and Kanzi as pupil, to demonstrate facility in language. It's also a high-stakes gambit. Because the obvious but unfortunate corollary is that if the case for real language skill is not persuasive, then the case for humane treatment fails with it, an outcome doubly tragic for Kanzi who spent his entire life in captivity to help decide the matter.

Whatever one's stance here on Kanzi, it clearly illustrates the elision of science and ethics. It highlights the appeal but also the challenges and potential risks associated with moving from descriptive science to normative science, from concerning ourselves with "*how things are*" to concerning ourselves also with "*how things ought to be.*" The danger, of course, is that our commitment to how we think things ought to be is likely to color how we then think things truly are: our ability to study and describe reality distorted by how we wish it were. Whatever the potential value of anthropomorphism, this is a serious potential pitfall.

It's important to be clear that, in cautioning against normative science and the logic in it that might put Kanzi's case for more compassionate and humane treatment on the block, or the same case for any other animal, there is absolutely no critique of their abilities, intelligence or inherent worth. On the contrary, we should celebrate the abilities of other species and accord them due compassion and protection – however, we should do that without requiring that their abilities also be our particular brand of ability. We should respect, support, protect and conserve them for their own sake and not for how much they remind us of ourselves, which is both anthropocentric and conceited, as though ours was the only kind of life worth valuing.

In fact, surely it is the more remarkable that bats navigate in the dark in the way that they do, with a sensory ability entirely foreign to us. Surely, it is awe-inspiring to think that they can see things much as we do – for they *can* also see – but that they can also effectively "see" acoustically, meaning that they also form auditory profiles of shapes and obstacles in the world around them. In essence, then, they know both what the landscapes they inhabit *look* like and also what they *sound* like, which is frankly a bit hard even to fathom. And surely that is the better part of the "grandeur in this view of life" to which Darwin so famously alluded.

CONCLUSION

Seyfarth et al. (1980a) published a landmark paper in *Science*, reporting evidence interpreted as language-like semantic communication in a primate. It had a profound impact. It represented the best case for semantic communication in animals and became the textbook example for it, catalyzing an enduring quest for semantics, syntax and other high-level features of human language in the communication of primates and other species. Ultimately, the original findings were revised, the case for vervet semantics disproven, and with it also support for the notion even of functional referentiality as the much diluted version of representation to which the primate semantics agenda had previously retreated (Wheeler and Fischer, 2012). Nevertheless, the broader research program engendered by the vervet story continues apace (Searcy, 2019). Why and how the vervet research had such impact warrants sober reflection.

There are a host of possible reasons why the vervet alarm story may have been so impactful originally, and why its legacy remains so enduring, a few of which have been considered here. They include the role of a contemporaneous paradigm shift in research on human and animal behavior and psychology in the middle 20th century and the role of key influencers in that shift. Chomsky was one key influencer in the Cognitive Revolution, following his mini-revolution in linguistics with bold proposals about Universal Grammar and an innate cognitive module dedicated to language acquisition [the so-called Language Acquisition Device (LAD)], both posited as necessary to overcome what he took to be the fundamental unlearnability of language by children. Griffin and Marler were key influencers carrying the cognitivist baton in the field of animal communication, where they followed Chomsky's lead in foregrounding human language as the benchmark for sophisticated cognition in animals and so used explicit comparisons to human language to frame comparative study of communication in other species.

The resulting cognitivist movement proved tremendously productive generally, including in comparative psychology and in research on animal communication. However, it has over time also confronted some of its own limits, as many of its founding findings and propositions have been eclipsed: Chomskyan Universal Grammar has been successively diluted and ultimately, largely abandoned; the need for an innate language acquisition module has been similarly obviated by research now demonstrating the fundamental learnability of language by children; modern cognitivism recapitulates core elements of behaviorism in associationist processes undergirding the contemporary focus on neural networks, connectionism, and deep learning; and recent reviews of primate communication now acknowledge it is not semantic nor meaningfully language-like. The disruptive influence of paradigm shifts can be extremely productive. However, invariably they also sow the seeds of their own succession. Paradigms change, that reality is built into the paradigm concept. The long history of reversals in variously embracing versus eschewing mentalistic constructs in scholarly discourse over the last few centuries should make us especially wary that the current cycle in this fashion is either the best or the last.

The original vervet alarm call story appealed during a particular cycle that spawned a thriving mentalist, cognitivist paradigm. That paradigm provided a welcoming intellectual climate prepared to receive the vervet findings. The work also had broader popular appeal, tapping an anthropomorphic instinct or habit, effectively pre-prepared to receive and reward reports of human-like traits in animals. Some researchers will bristle at the latter suggestion. They will reject the possibility that anthropomorphism ever influenced their own work, being acutely aware of its checkered reputation in scholarly circles if not aware of its possible unconscious influence on them nevertheless. After all, instinct and habit are difficult to control. That's the point. They are habitual and instinctual precisely because they are not under explicit and reflective control: like other common biases we continue to wrestle with – gender bias, racial bias – that operate widely, unconsciously, unwittingly. We are often not even aware of their influence. And even when we are aware, they can remain (maddeningly) immune to our formal, conscious and explicit attempts to mitigate and eliminate their influence.

For some, this history and its outcomes will give pause. That the original claims of semanticity in vervets have been recanted, and with them also broader claims of semantic communication in primates, may be troubling, possibly creating doubts about much else besides. Others may not be so affected, confident in the value of the much wider program of research on language parallels in animals that the vervet work helped to catalyze. There are many possible conclusions that could be drawn, and no doubt many possible opinions about the importance of other factors at play – formal intellectual paradigm shifts, spurred by key influencers, intersecting other universal human biases. But whatever one's predilection on these matters, perhaps there can, at the least, be common commitment to the impartial and timeless prescription of Laplace, Williams and Sagan that our production and reception of research maintain a healthy circumspection; that we value and expect high standards of evidence, perhaps particularly for claims of an extraordinary nature.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The author acknowledges funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada (Grant #: 217307-2013) and general research support from the University of New Brunswick.

ACKNOWLEDGMENTS

The author acknowledges and is grateful for continuing research support from the Natural Sciences and Engineering Research Council (NSERC) of Canada, as well as the University of New Brunswick. I also acknowledge that my wife and I have a cat, fed fresh shrimp twice daily. We talk to him all the time but are pretty certain he doesn't understand us. We're also pretty certain he doesn't actually care to Gray (2020). That's part of the grandeur of cats, and something else to be grateful for. This article is dedicated to long-time collaborator and friend, Michael J. Owren, and his cat, Mama Cass, as well as Harley, Tucker, OJ, and Skookum.

REFERENCES

Arnold, K., and Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. *Curr. Biol.* 18, R202–R203.

Barrett, H. C. (2005). "Cognitive development and the understanding of animal behavior," in *Origins of the Social Mind*, eds B. J. Elis and D. F. Bjorklund (New York, NY: The Guilford Press), 438–467.

Barrett, L., Henzi, S. P., and Rendall, D. (2007). Social brains, simple minds: does social complexity really require cognitive complexity? *Philos. Trans. R. Soc. B* 362, 561–575. doi: 10.1098/rstb.2006.1995

Berwick, R. C., and Chomsky, N. (2016). *Why Only Us: Language and Evolution.* Cambridge, MA: MIT Press.

Bickerton, D. (1992). *Language and Species*. Chicago: University of Chicago Press.

Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., et al. (2017). Evidence of a vocalic proto-system in the baboon (*Papio Papio*) suggests pre-hominin speech precursors. *PLoS One* 12:e0169321. doi: 10.1371/journal.pone.0169321

Brentano, F. (1874). *Psychology from an Empirical Standpoint*. London: Routledge and Kegan Paul.

Budiansky, S. (1998). *If a Lion Could Talk: Animal Intelligence and the Evolution of Consciousness*. New York, NY: The Free Press.

Call, J., and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192. doi: 10.1016/j.tics.2008.02.010

Cheney, D. L., and Seyfarth, R. M. (1990a). Attending to behavior versus attending to knowledge: examining monkeys' attribution of mental states. *Anim. Behav.* 40, 742–753. doi: 10.1016/s0003-3472(05)80703-1

Cheney, D. L., and Seyfarth, R. M. (1990b). *How Monkeys See the World: Inside the Mind of Another Species*. Chicago: University of Chicago Press.

Cheney, D. L., and Seyfarth, R. M. (1998). "Why animals don't have language," in *The Tanner Lectures on Human Values*, ed. G. B. Pearson (Salt Lake City, UT: University of Utah Press), 174–209.

Cheney, D. L., and Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *Linguist. Rev.* 22, 135–159.

Cheney, D. L., and Seyfarth, R. M. (2008). *Baboon Metaphysics: The Evolution of the Social Mind*. Chicago: University of Chicago Press.

Cheney, D. L., Seyfarth, R. M., and Palombit, R. A. (1996). The function and mechanisms underlying baboon 'contact' barks. *Anim. Behav.* 52, 507–518. doi: 10.1006/anbe.1996.0193

Crockford, C., Wittig, R. M., and Zuberbühler, K. (2015). An intentional vocalization draws others' attention: a playback experiment with wild chimpanzees. *Anim. Cogn.* 18, 581–591. doi: 10.1007/s10071-014-0827-z

Davila Ross, M., Menzler, S., and Zimmerman, E. (2008). Rapid facial mimicry in orangutan play. *Biol. Lett.* 4, 27–30. doi: 10.1098/rsbl.2007.0535

Davila Ross, M., Owren, M. J., and Zimmerman, E. (2010). The evolution of laughter in great apes and humans. *Commun. Integr. Biol.* 3, 191–194. doi: 10.4161/cib.3.2.10944

de Saussure, F. (1971). *Cours de Linguistique Générale*. Paris: Payot. Originally published in 1916.

de Waal, F. B. M. (1999). Anthropomorphism and anthropodenial: Consistency in our thinking about humans and other animals. *Philos. Top.* 27, 255–280. doi: 10.5840/philtopics199927122

de Waal, F. B. M. (2018). *Mama's Last Hug: Animal Emotions and What They Tell Us About Ourselves*. New York, NY: W.W. Norton & Company.

Deacon, T. (1998). *The Symbolic Species*. New York, NY: W.W. Norton & Company.

Dennett, D. (1983). Intentional systems in cognitive ethology: the "Panglossian paradigm" defended. *Behav. Brain Sci.* 6, 343–390. doi: 10.1017/s0140525x00016393

Dennett, D. (1995). *Darwin's Dangerous Idea*. New York, NY: Simon & Schuster.

Dezecache, G., Zuberbühler, K., Davila-Ross, M., and Dahl, C. D. (2020). Flexibility in wild infant chimpanzee vocal behavior. *J. Lang. Evol.* 6, 37–53. doi: 10.1093/jole/lzaa009

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol.* 6, 178–190.

Eddy, T. J., Gallup, G. G., and Povinelli, D. J. (1993). Attribution of cognitive states to animals: anthropomorphism in comparative perspective. *J. Soc. Issues* 49:87e101.

Evans, C. S. (1997). "Referential signals," in *Perspectives in Ethology: Communication*, Vol. 12, eds D. H. Owings, M. D. Beecher, and N. S. Thompson (New York, NY: Plenum), 99–143. doi: 10.1007/978-1-4899-1745-4_5

Evans, C. S., Evans, L., and Marler, P. (1993). On the meaning of alarm calls: functional reference in an avian vocal system. *Anim. Behav.* 46, 23–38. doi: 10.1006/anbe.1993.1158

Fischer, J. (2021). Primate vocal communication and the evolution of human speech. *Curr. Direct. Psychol. Sci.* 83, 153–173. doi: 10.3378/027.083.0202

Fischer, J., and Price, T. (2017). Meaning, intention and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* 82, 22–31. doi: 10.1016/j.neubiorev.2016.10.014

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends Cogn. Sci.* 4, 258–267. doi: 10.1016/s1364-6613(00)01494-7

Fitch, W. T. (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.

Fitch, W. T. (2020). Animal cognition and the evolution of language: why we cannot focus solely on communication. *Philos. Trans. R. Soc. B* 375:20190046. doi: 10.1098/rstb.2019.0046

Ghazanfar, A. A., Liao, D. A., and Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Anim. Behav.* 15, 239–247. doi: 10.1016/j.anbehav.2019.01.021

Ghazanfar, A. A., and Rendall, D. (2008). The evolution of human vocal production. *Curr. Biol.* 18, R457–R460.

Gillispie, C. C. (2000). *Pierre Simon Laplace: A Life in Exact Science*. Princetion, NJ: Princeton University Press.

Gouzoules, S., Gouzoules, H., and Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: Representational signalling in the recruitment of agonistic aid. *Anim. Behav.* 32, 182–193. doi: 10.1016/s0003-3472(84)80336-x

Gray, J. (2020). *Feline Philosophy*. New York, NY: Farra, Strauss & Giroux.

Green, S. (1975). "Variation of vocal pattern with social situation in the Japanese monkey (Macaca fuscata): a field study," in *Primate Behaviour*, Vol. 4, ed. L. A. Rosenblum (New York, NY: Academic Press), 1–102. doi: 10.1016/b978-0-12-534004-5.50006-3

Greene, E., and Meagher, T. (1998). Red squirrels produce predator-class specific alarm calls. *Anim. Behav.* 55, 511–518. doi: 10.1006/anbe.1997.0620

Grice, H. P. (1957). Meaning. *Philos. Rev.* 66, 377–388.

Griebel, U., Pepperberg, I. M., and Oller, D. K. (2016). Developmental plasticity and language: a comparative perspective. *Top. Cogn. Sci.* 8, 435–445. doi: 10.1111/tops.12200

Griffin, D. (1985). *Animal Thinking*. Cambridge, MA: Harvard University Press .

Griffin, D. R. (1976). *The Question of Animal Awareness*. New York, NY: Rockefeller University Press.

Griffin, D. R. (1995). Windows on animal minds. *Conscious. Cogn.* 4, 194–204. doi: 10.1006/ccog.1995.1025

Hage, S. R. (2010). "Neuronal networks involved in the generation of vocalization," in *Handbook of Behavioral Neuroscience*, ed. S. Brudzynski (Oxford: Elsevier), 339–349. doi: 10.1016/b978-0-12-374593-4.00032-2

Hammerschmidt, K., and Fischer, J. (2008). "Constraints in primate vocal production," in *The Evolution of Communicative Creativity: From Fixed Signals to Contextual Flexibility*, eds K. Oller and U. Griebel (Cambridge, MA: The MIT Press), 93–119.

Hardy, B. L., Moncel, M.-H., Kerfant, C., Leblon, M., Bellot-Gurlet, L., and Melard, N. (2020). Direct evidence of Neanderthal fibre technology and its cognitive and behavioral implications. *Sci. Rep.* 10:4889.

Hauser, M. D. (1998). Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim. Behav.* 55, 1647–1658. doi: 10.1006/anbe.1997.0712

Hauser, M. D. (2009). The possibility of impossible cultures. *Nature* 460, 190–196. doi: 10.1038/460190a

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., et al. (2014). The mystery of language evolution. *Front. Psychol.* 5:401. doi: 10.3389/fpsyg.2014.00401

Hodos, W., and Campbell, C. B. G. (1969). Scala naturae: why there is no theory in comparative psychology. *Psychol. Rev.* 76, 337–350. doi: 10.1037/h0027523

Hodos, W., and Campbell, C. B. G. (1990). "Evolutionary scales and comparative studies of animal cognition," in *Neurobiology of Comparative Cognition*, eds R. P. Kesner and D. S. Olton (Hillsdale, NJ: Lawrence Erlbaum Associates), 1–20.

Humphrey, N. (1976). "The social function of intellect," in *Growing Points in Ethology*, eds P. P. G. Bateson and R. A. Hinde (Cambridge, MA: Cambridge University Press), 303–317.

Jolly, A. (1966). Lemur social behaviour and primate intelligence. *Science* 153, 501–506. doi: 10.1126/science.153.3735.501

Jürgens, U. (2008). The neural control of vocalization in mammals: a review. *J. Voice* 23, 1–10.

Kennedy, J. S. (1992). *The New Anthropomorphism*. Cambridge: Cambridge University Press.

Kitzmann, C., and Caine, N. (2009). Marmoset (*Callithrix geoffroyi*) food-associated calls are functionally referential. *Ethology* 115, 439–448. doi: 10.1111/j.1439-0310.2009.01622.x

Klopfer, P. (2005). Animal cognition and the new anthropomorphism. *Intern. J. Comp. Psychol*. 18, 202–206.

Krebs, J. R., and Dawkins, R. (1984). "Animal signals: mind-reading and manipulation," in *Behavioural Ecology: An Evolutionary Approach*, 2nd Edn, eds J. R. Krebs and N. B. Davies (Oxford: Blackwell Scientific), 380–402.

Leakey, R., and Lewin, R. (1977). *Origins: What New Discoveries Reveal About the Emergence of our Species and Its Possible Future*. New York, NY: E.P. Dutton.

Leger, D. W. (1993). Contextual sources of information and responses to animal communication signals. *Psychol. Bull.* 113, 295–304. doi: 10.1037/0033-2909.113.2.295

Locke, J. (2021). The indexical voice: communication of personal states and traits in humans and other primates. *Front. Psychol.* 12:651108. doi: 10.3389/fpsyg.2021.651108

Macedonia, J. M., and Evans, C. S. (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* 93, 177–197. doi: 10.1111/j.1439-0310.1993.tb00988.x

Manser, M. B., Seyfarth, R. L., and Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *Trends Cogn. Sci.* 6, 55–57. doi: 10.1016/s1364-6613(00)01840-4

Marler, P. (1977). "Primate vocalization: affective or symbolic?," in *Progress in Great Ape Research*, ed. G. H. Bourne (New York, NY: Academic Press).

Marler, P., Evans, C. S., and Hauser, M. D. (1992). "Animal signals: motivational, referential, or both?," in *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, eds H. Papousek, U. Juürgens , and M. Papousek (Cambridge: Cambridge University), 66–86.

Mitani, J. C., Call, J., Kappeler, P. M., Palombit, R. A., and Silk, J. B. (2012). *The Evolution of Primate Societies*. Chicago: University of Chicago Press.

Nieder, A., and Mooney, R. (2020). The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philos. Trans. R. Soc. B* 375:20190054. doi: 10.1098/rstb.2019.0054

Nielsen, A., and Rendall, D. (2018). "Comparative perspectives on communication in human and nonhuman primates: grounding meaning in shared processes of vocal production, perception, affect and cognition," in *Oxford Handbook of Voice Perception*, eds S. Frühholz and P. Belin (Oxford: Oxford University Press), 279–308.

Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., et al. (2019). Language origin seen in spontaneous and interactive vocal rate of human and bonobo infants. *Front. Psychol.* 10:729. doi: 10.3389/fpsyg.2019.00729

Owings, D. H. (1994). How monkeys feel about the world: a review of how monkeys see the world. *Lang. Commun.* 14, 15–30. doi: 10.1016/0271-5309(94)90017-5

Owings, D. H., and Morton, E. S. (1998). *Animal Vocal Communication: A New Approach*. Cambridge: Cambridge University Press.

Owren, M. J., and Rendall, D. (2001). Sound on the rebound: returning form and function to the forefront in understanding nonhuman primate vocal signaling. *Evol. Anthropol.* 10, 58–71. doi: 10.1002/evan.1014

Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: Influence versus information in communication. *Biol. Philos.* 25, 755–780. doi: 10.1007/s10539-010-9224-4

Penn, D. C., and Povinelli, D. J. (2007). On the lack of evidence that chimpanzees possess anything remotely resembling a 'theory of mind.'. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 362, 731–744. doi: 10.1098/rstb.2006.2023

Pinker, S. (1994). *The Language Instinct*. New York, NY: William Morrow.

Pinker, S., and Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition* 95, 201–236.

Pomberger, T., Risueno-Segovia, C., Loschner, J., and Hage, S. R. (2018). Precise motor control enables rapid flexibility in vocal behavior of marmoset monkeys. *Curr. Biol.* 28, 788–794. doi: 10.1016/j.cub.2018.01.070

Povinelli, D., Bering, J., and Giambrone, S. (2000). Toward a science of other minds: escaping the argument by analogy. *Cogn. Sci.* 24, 509–541. doi: 10.1207/s15516709cog2403_7

Premack, D. (1972). Concordant preferences as a precondition for affective but not for symbolic communication (or How to do experimental anthropology). *Cognition* 1, 251–264. doi: 10.1016/0010-0277(72)90021-2

Premack, D. (1985). 'Gavagai!' or the future history of the animal language controversy. *Cognition* 19, 207–296. doi: 10.1016/0010-0277(85)90036-8

Price, T. (2013). *Vocal Communication within the Genus Chlorocebus: Insights into Mechansisms of Call Production and Call Perception*. Ph. D thesis. Göttingen: University of Gottingen.

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5, 1–11.

Radick, G. (2007). *The Simian Tongue: The Long Debate about Animal Language*. Chicago: University of Chicago Press.

Rendall, D., Cheney, D. L., and Seyfarth, R. M. (2000). Proximate factors mediating 'contact' calls in adult female baboons and their infants. *J. Compar. Psychol.* 114, 36–46. doi: 10.1037/0735-7036.114.1.36

Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (F_0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *J. Acoustic. Soc. Am.* 117, 944–955. doi: 10.1121/1.1848011

Rendall, D., Notman, H., and Vokey, J. R. (2007). "Homologizing the mind: Frodoian or Fodorian psychology," in *Handbook of Evolutionary Psychology*, eds L. Barrett and R. I. M. Dunbar (Oxford: Oxford University Press), 59–70.

Rendall, D., and Owren, M. J. (2013). "Communication without meaning or information: Abandoning language-based and informational constructs in animal communication theory," in *Animal Communication Theory: Information and Influence*, ed. U. Stegmann (Cambridge: Cambridge University Press), 151–182. doi: 10.1017/cbo9781139003551.010

Rendall, D., Owren, M. J., and Rodman, P. S. (1998). The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *J. Acoustic. Soc. Am.* 103, 602–614. doi: 10.1121/1.421104

Rendall, D., Owren, M. J., and Ryan, M. J. (2009). What do animal signals mean? *Anim. Behav.* 78, 233–240. doi: 10.1016/j.anbehav.2009.06.007

Rendall, D., Owren, M. J., Weerts, E., and Hienz, R. D. (2004). Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. *J. Acoustic. Soc. Am.* 115, 411–421. doi: 10.1121/1.1635838

Romanes, G. J. (1882). *Animal Intelligence*. Boston, MA: D. Appleton and Company.

Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., et al. (2014). The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507, 354–357. doi: 10.1038/nature12961

Savage-Rumbaugh, S., and Lewin, R. (1994). *Kanzi: The Ape at the Brink of the Human Mind*. New York, NY: Wiley.

Savage-Rumbaugh, S., Wamba, K., Wamba, P., and Wamba, N. (2007). Welfare of apes in captive environments: comments on, and by, a specific group of apes. *J. Appl. Anim. Welf. Sci.* 10, 7–19. doi: 10.1080/10888700701277261

Scott-Phillips, T. C. (2015). Meaning in animal and human communication. *Anim. Cogn.* 18, 801–805. doi: 10.1007/s10071-015-0845-5

Searcy, W. A. (2019). Animal communication, cognition, and the evolution of language. *Anim. Behav.* 151, 203–205. doi: 10.1016/j.anbehav.2019.03.001

Seyfarth, R. M., and Cheney, D. L. (1984). Grooming, alliances, and reciprocal altruism in vervet monkeys. *Nature* 308, 541–543. doi: 10.1038/308541a0

Seyfarth, R. M., and Cheney, D. L. (2017). The origin of meaning in animal signals. *Anim. Behav.* 124, 339–346. doi: 10.1016/j.anbehav.2016.05.020

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980a). Monkey responses to three different alarm calls: evidence for predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980b). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28, 1070–1094. doi: 10.1016/s0003-3472(80)80097-2

Slobodchikoff, C. N., Paseka, A., and Verdolin, J. L. (2009). Prairie dog alarm calls encode labels about predator colors. *Anim. Cogn.* 12, 435–439. doi: 10.1007/s10071-008-0203-y

Slocombe, K. E., and Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proc. Natl. Acad. Sci. U.S.A.* 104:17228. doi: 10.1073/pnas.0706741104

Smith, W. J. (1977). *The Behavior of Communicating*. Cambridge, MA: Harvard University Press.

Smuts, B. B., Cheney, D. L., Seyfarth, R. M., and Wrangham, R. W. (1988). *Primate Societies*. Chicago: University of Chicago Press.

Sober, E. (1998). "Morgan's canon," in *The Evolution of Mind*, eds C. Allen and D. Cummins (Oxford: Oxford University Press), 224–242.

Soltis, J. (2004). The functions of early infant crying. *Behav. Brain Sci.* 27, 443–490. doi: 10.1017/s0140525x0400010x

Struhsaker, T. T. (1967). *Behavior of Vervet Monkeys (Cercopithecus aethiops)*. Berkeley, CA: University of California Press.

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., et al. (2015). The developmental dynamics of marmoset monkey vocal production. *Science* 6249, 734–738. doi: 10.1126/science.aab1058

Templeton, C. N., Greene, E., and Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science* 308, 1934–1937. doi: 10.1126/science.1108841

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Urquiza-Haas, E. G., and Kotrschal, K. (2015). The mind behind anthropomorphic thinking: attribution of mental states to other species. *Anim. Behav.* 109, 167–176. doi: 10.1016/j.anbehav.2015.08.011

von Uexkuüll, J. (1909). *Umwelt und Innenwelt der Tiere*. Berlin: Verlang von Julius Springer.

Wheeler, B. C., and Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evol. Anthropol.* 21, 195–205. doi: 10.1002/evan.21319

Williams, G. C. (1966). *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton, NJ: Princeton University Press.

Wynne, C. (2004). The perils of anthropomorphism. *Nature* 428:606. doi: 10.1038/428606a

Zeifman, D. M. (2001). An ethological analysis of human infant crying: answering Tinbergen's four questions. *Dev. Psychobiol.* 39, 265–285. doi: 10.1002/dev.1005

Zoloth, S., and Green, S. (1979). Monkey vocalizations and human speech: Parallels in perception? *Brain Behav. Evol.* 16, 430–442. doi: 10.1159/0001 21880

Zuberbühler, K. (2019). Evolutionary roads to syntax. *Anim. Behav.* 151, 259–265. doi: 10.1016/j.anbehav.2019.03.006

Zuberbühler, K., Cheney, D. L., and Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *J. Compar. Psychol.* 113, 33–42. doi: 10.1037/0735-7036.113.1.33

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 23 September 2021
doi: 10.3389/fpsyg.2021.647841

Nonhuman and Nonhuman-Human Communication: Some Issues and Questions

*Irene M. Pepperberg**

Department of Psychology, Harvard University, Cambridge, MA, United States

Deciphering nonhuman communication – particularly nonhuman vocal communication – has been a longstanding human quest. We are, for example, fascinated by the songs of birds and whales, the grunts of apes, the barks of dogs, and the croaks of frogs; we wonder about their potential meaning and their relationship to human language. Do these utterances express little more than emotional states, or do they convey actual bits and bytes of concrete information? Humans' numerous attempts to decipher nonhuman systems have, however, progressed slowly. We still wonder why only a small number of species are capable of vocal learning, a trait that, because it allows for innovation and adaptation, would seem to be a prerequisite for most language-like abilities. Humans have also attempted to teach nonhumans elements of our system, using both vocal and nonvocal systems. The rationale for such training is that the extent of success in instilling symbolic reference provides some evidence for, at the very least, the cognitive underpinnings of parallels between human and nonhuman communication systems. However, separating acquisition of reference from simple object-label association is not a simple matter, as reference begins with such associations, and the point at which true reference emerges is not always obvious. I begin by discussing these points and questions, predominantly from the viewpoint of someone studying avian abilities. I end by examining the question posed by Premack: do nonhumans that have achieved some level of symbolic reference then process information differently from those that have not? I suggest the answer is likely "yes," giving examples from my research on Grey parrots (*Psittacus erithacus*).

Keywords: interspecies communication, symbolic reference, nonhuman communication, animal cognition, Grey parrot cognition

OPEN ACCESS

Edited by:
David M. G. Lewis,
Murdoch University, Australia

Reviewed by:
Erich David Jarvis,
Duke University, United States
Eduardo Mercado,
University at Buffalo, United States

****Correspondence:***
Irene M. Pepperberg
impepper@media.mit.edu

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *30 December 2020*
Accepted: *25 August 2021*
Published: *23 September 2021*

Citation:
Pepperberg IM (2021) Nonhuman and Nonhuman-Human Communication: Some Issues and Questions.
Front. Psychol. 12:647841.
doi: 10.3389/fpsyg.2021.647841

INTRODUCTION

The songs of whales and birds, the roars of lions and bellows of elephants, the pant-hoots and grunts of apes, the squeaks of mice and croaks of frogs – humans have long been fascinated by the meanings of nonhuman communications systems. And those examples involve only the auditory mode – what about the flashes of lightning bugs or the scent systems that make dogs stop and sniff every few feet on their daily walk? Lest any doubt exist about the longstanding and widespread nature of such interest amongst even nonscientists, one need only cite examples such as the legend of King Solomon's ring, which purportedly enabled him to communicate at will with all the birds and beasts in his realm (Lorenz, 1952); the historic lore of Native Americans, who supposedly could change into various animals and thus share their lives (Rasmussen, 1972), or the many children's books on the subject (e.g., the Dr. Doolittle series; Lofting, 1920). Darwin (1871) in particular wrote at length about the similarities between human and nonhuman communication systems but provided scant guidance for deciphering the latter. Realistically, in his era little experimental research was possible that could have discovered potential meaning in nonhuman systems or their relationship to human language: no one yet had the appropriate tools to determine whether nonhuman signals expressed anything more than emotional states or conveyed actual bits and bytes of concrete information.

Such research began in earnest in the first half of the 1900s but was extremely limited in scope, as the methodology was hardly more advanced than in Darwin's day. Nice (1943) and Saunders (1951), for example, were among the first to quantify and contextualize birdsong in a scientific manner and demonstrate the inherent complexity in various species' systems. They employed musical notation and detailed field notes to describe the vocalizations of birds and various contexts in which such vocalizations were used in species such as song sparrows. Researchers, such as Marler (1956) and Thorpe (1958), continued this work with a variety of species, and pioneered use of the Sonagraph©, which gave plots of time vs. frequency (kHz) that enabled these songs to be analyzed in a myriad of ways. However, these researchers were a small minority among scientists who were, particularly in the early half of the 20th century, more interested in collecting skins and categorizing species than in studying behavior – especially communication.

Not until the latter half of that century would tools and techniques be designed that allowed humans to engage in the detailed analyses and formal experiments that would provide the first real insights into the realm of animal communication systems, primarily with respect to birdsong. High-quality microphones and tape recorders ("portable" only in the sense that they could be transported into the field!) allowed researchers to acquire recordings of actual songs as they were sung in nature; the Sonagraph© and these recorders also enabled researchers to choose and play specific songs back to the birds to see what effects hearing these had on behavior. And, thus began the study of nonhuman communication systems in earnest…

BIRDSONG, PRIMATE CALLS, AND VOCAL LEARNING

The second half of the 20th century was an especially exciting time for the study of nonhuman behavior. In 1973, for the first time ever, three ethologists (Nikolaas Tinbergen, Konrad Lorenz, and Karl von Frisch) won the Nobel Prize in Physiology or Medicine, a solid recognition of the validity of the field. Notably, the work of each of them involved, to greater or lesser degrees, nonhuman communication. From the standpoint of someone like myself, interested in vocal interactions (particularly in birds), the wealth of new information that was beginning to accrue was almost overwhelming. Although earlier researchers (e.g., Howard, 1920) had proposed that song had specific purposes – territorial defense and mate attraction – they lacked clear evidence to support these claims. Now researchers were able to acquire data about the actual meaning and function of avian signals (Pepperberg, 2020c): Dilger (1956) showed that thrushes used song to distinguish subspecies from one another as well as to defend their territories from competitors. Marler (1952) discovered local dialects in the songs of chaffinches, which steered Thorpe (1958) to the study of their song acquisition and led both scientists to the investigation of vocal learning – if song was innately specified, all birds of the same species should sound almost exactly the same; thus, the existence of dialects suggested that differences were acquired. (I will discuss song learning in detail later.). Weeden and Falls (1959) revealed that ovenbirds used their songs for individual recognition; Konishi (1964), for example, found that the trill part of the song served that purpose in Oregon juncos. Many researchers (e.g., Krebs et al., 1978; Yasukawa, 1981) used playbacks to determine how individual recognition was used to discriminate neighboring males (those birds who had pretty much defined their abutting territories and who used countersinging to keep the status quo) from stranger males (those who were not local and posed a serious disruptive threat to the status quo). Marler (1960) provided some of the earliest evidence that songs were also used for mate attraction and was supported by somewhat later studies such as those by Kroodsma (1976) and Krebs (1977) on multiple uses of song, and how different songs could be used for different purposes (e.g., see Catchpole, 1983; later Kroodsma et al., 1989). Marler (1961) and Smith (1963) separately systematized the analyses of nonhuman communication systems, drawing heavily on avian studies. Studies in some non-songbirds – parrots – showed that these birds also exhibited individual vocal recognition and alarm calling, suggestive of sentinel behavior, that alerted the flock to various flying predators (Lawson and Lanning, 1980; Levinson, 1980).

Note that many of the same techniques were being used to study vocal behavior in nonhuman primates (e.g., Struhsaker, 1967; Seyfarth et al., 1980a,b), which led to the claim that vervet monkeys had referential, vocal signaling – different calls for different types of predators (i.e., their argument was that each signal denoted – referred to – a specific entity; I will discuss the meaning of "reference" and various levels of referential behavior more fully in a separate section). The vervet study stirred up considerable controversy as to the extent or level of the referentiality exhibited (Pepperberg, 2020c): Nay-sayers argued that, unlike humans, vervets could not "discuss" predators outside of alarm calling in their presence (see Bickerton, 1990), and even Marler (1974) noted that the calls could simply be context-dependent (i.e., merely indicating "danger above" vs. "danger below" instead of referring to specific predators). Note that the controversy continues (see Fischer, 2011; Wheeler and Fischer, 2012; Townsend and Manser, 2013), even Seyfarth and Cheney eventually backed off somewhat in the strength of their claims (see Price et al., 2014). Whatever the level of reference, however, the calls of both birds and primates did contain *content* – warning signals. Referentiality, as I will discuss later, is an especially fraught topic.

Despite advances in the realm of nonhuman primate communication, the study of birdsong retained a very special status in the research community for one interesting reason: Marler (1970a) and Thorpe (1974), based on pioneering studies by researchers like Nottebohm (1966, 1970), had found that many avian species – like humans and *unlike* nonhuman primates – were vocal learners and, like humans, have lateralized brain areas responsible for such acquisition. These researchers therefore suggested that birds could be better models than nonhuman primates for studying the evolution of language, despite being more evolutionarily distant from humans than nonhuman primates. Specifically, vocal learning allows for innovation and adaptation, characteristics that release communication from rigid, innately specified responses to stimuli. Although no one then (or even now) could claim that birds' songs have the kind of syntactical and semantic complexity of human language, researchers did demonstrate numerous parallels between song and language (a brief summary follows; for full reviews see, for example, Doupe and Kuhl, 1999; Peters and Nowicki, 2017; also, see Beecher, in this collection of papers).

Starting with the earliest stages of acquisition, both birds (Marler, 1970a) and humans (e.g., Oller et al., 1976) engage in a period in which they "babble" – that is, they experiment with the sounds that will ultimately become the building blocks of their repertoire. Moreover, in both cases, the babbling itself consists of stages, such that progress toward the adult behavior can be monitored (e.g., Marler, 1970b; Oller, 1978; de Boysson-Bardies et al., 1989). Of particular interest is that both birds and humans initially produce an extremely wide variety of sounds – formally termed *overproduction* – of which only a subset will eventually be used in their communication systems. This initial overproduction enables the possibility, respectively, of producing many songs/dialects and all human languages. Over the course of maturation, a winnowing occurs, based on cultural input, to focus learning on the most relevant sounds in their communicative environment (e.g., Rice and Thompson, 1968; Doupe and Kuhl, 1999).

Another common aspect involves the issue of what were initially called "critical periods" for acquisition – limited, tightly time-specified windows during which exposure to the adult system was considered necessary if learning was to occur (e.g., Lenneberg, 1967; Marler, 1970a). Such windows are now considered to be much less time-constrained, given that subsequent research has demonstrated how, for example, exposure to intense, live, social interactions rather than auditory tapes can greatly extend the period for acquisition (e.g., Fromkin et al., 1974; Baptista and Petrinovich, 1984, 1986; Grimshaw et al., 1998; Beecher and Brenowitz, 2005). Thus, the preferred phrase to describe these intervals is now "sensitive phases." Nevertheless, for both humans and birds, early exposure allows the most *facile* development. Of note, however, are those avian species such as canaries, starlings, and parrots that are considered "open-ended learners" (Adret-Hausberger et al., 1990; Nottebohm, 2006); these birds have the capacity to acquire new vocalizations throughout their lives. A related aspect involves the issue of bilingualism – in humans, the acquisition of a second language; in birds, the acquisition of a second dialect or even the song of another species (allospecific song). Early studies on birds, again using only auditory tapes (Marler, 1970a), argued that allospecific song acquisition was prevented by an innate template, and also argued that the same template likely prevented acquisition of multiple dialects; early studies on humans argued that bilingualism was limited by the same critical period that was thought to constrain first language (Johnson and Newport, 1989). Again, subsequent studies demonstrated that, depending upon the type and extent of input, later acquisition can indeed occur – that the length of the sensitive phase can be extended, quite significantly, in both birds and humans (e.g., Baptista and Petrinovich, 1984, 1986; Hartshorne et al., 2018).

The use of birds as models for language evolution was made even stronger by the studies on the neurobiology and neuroanatomy of the vocal learning system in birds and humans. For a detailed but still concise review, see Jarvis (2019). Briefly, although the capacity for vocal learning most likely arose independently in birds (parrots, songbirds, and hummingbirds) and mammals (which include bats, elephants, and marine mammals as well as humans), specific analogous brain areas and connections that are responsible for that ability have been found in humans and avian vocal learners (note Colquitt et al., 2021) and particularly in vocal open-ended learners; these areas and connections are lacking in closely-related species that are not vocal learners. According to Jarvis (2019, p. 4):

> "the common ancestor of vertebrates had a brainstem pathway for production of innate vocalizations with limited vocal plasticity… In some species, the forebrain motor learning pathway then duplicated and formed a vocal motor learning pathway with weak direct projections to the brainstem vocal motor neurons. Thereafter, this forebrain vocal motor learning pathway expanded in neuron numbers causing greater density of neurons in the forebrain, moved outside of the motor learning pathway, and gained dense direct projections to brainstem vocal motor neurons. Finally, the vocal learning pathway then duplicated one or more times and took on additional specialized gene regulation and connections, resulting in the advanced vocal learning pathways we find in parrots and in humans."

Although none of this detailed neurobiological information was known in the 1970s, enough about the avian brain had been learned (e.g., Nottebohm, 1970) to serve as the basis for the extensive research on avian vocalizations that led to these discoveries. The main points, of course, are that (a) for three avian lineages, there exist hundreds (parrots and hummingbirds) to thousands (oscines) of species in which vocalizations are learned, (b) learning is possible because of specific neuroanatomical areas and their functions, and (c) for those and many more species, vocalizations do contain information that (based on the actions of the receiver) appears to be processed in meaningful ways, even if humans have so far been unable to decipher much beyond those related to territorial defense, threats, mating and, in later studies, how the use of different songs or singing patterns in different contexts relates to degrees of aggressive behavior (see, for example, Smith, 1996; Beecher et al., 2000). I will henceforth use the term "meaningful" to define the vocalizations described in (c) above. I am not arguing that the behavior of the receiver is simply (possibly in a stimulus–response manner) triggered by the signal or manipulated by the signaler, but rather that the receiver *processes* (i.e., actively decodes and then integrates into prior and current knowledge) the information in the signal (plus possible other relevant input) and then decides on an appropriate course of action (note Seyfarth et al., 2010 for a discussion of this point).

AVIAN-HUMAN DIFFERENCES

Despite all these exciting parallels, some striking differences were found to exist between avian and human communication systems. Bird vocalizations demonstrate levels of *functional* reference (i.e., alarm calls, songs for mate attraction vs. territorial defense all encode *relatively* specific information about eliciting events, but must be processed with respect to the full context in which they are emitted); however, they apparently lack the kind of semanticity and syntax found in human speech. Specifically, Marler (1977) has characterized birdsong as having only a "phonological syntax"; that is, as a system in which the acoustic elements are arranged according to set rules in order to attract mates, deter rivals, and defend territory. One can argue that such is also true of combinations of physical and vocal displays used to extend the meaning of songs in a repertoire, especially when a bird has only a single song (e.g., Smith, 1996). However, the content of any specific *element* (for example, a note or syllable) does not (at least to our current knowledge) appear meaningful (i.e., in the sense of a human word; see below for additional information). Furthermore, such ordering is not a requisite for all avian species: For some vocal learners, note and syllable order is crucial for meaning and function; for other species it is not (see review in Weisman and Ratcliffe, 1987). Note that a more appropriate characterization of the latter might be that of *phonetic patterning* (Collier et al., 2014). However, for species in the former group (e.g., chestnut-sided warblers, *Dendroica pensylvanica*, Byers, 1995), particular songs – which consist of the same, but re-ordered, acoustic elements – *do* have different functions. Such distinctions may also be the case for certain bird calls – in particular instances, when the order of the elements is altered, birds fail to respond in playback tests (Suzuki et al., 2019). These data suggest that some sorts of rules for the production and comprehension of vocalizations may exist for some avian species, even though they are separated from humans by over 300 million years of evolution (Hedges et al., 1996). Human language, however, depends on the semantic meaning of each element of each sentence, as well as its hierarchical structure (e.g., Berwick et al., 2012). Moreover, elements of birdsong or calls are not (again, as far as humans have been able to determine at present!) routinely combined to form *novel* meanings for *novel* situations, as are human utterances. The only evidence for such avian combinations involves single behavioral instances and not specific individual vocal elements; for example, a Florida scrub jay (*Aphelocoma coerulescens*) once combined alarm calls associated with, respectively, hawks and snakes in the presence of a perched owl (Hailman and Elowson, 1984).

Thus, despite all these parallels between human language and birdsong, and arguments about nonhuman primate calls, researchers still were unable to determine the *extent* to which nonhuman communication was meaningful. To reiterate, experiments had shown that nonhuman signals *did* convey a certain amount of content – upon hearing certain signals, receivers knew what species was vocalizing, whether they should take evasive action from a predator and often what kind of action, whether another individual was trying to take over their territory or simply maintain a mutual boundary, whether a male was still searching for a mate – but this content involved basic behavioral states and concrete information. What seemed to be lacking was the type of *abstract* meaning that characterized human signals; for example, the ability to describe or comment upon something that was not physically present or is imaginary or the ability to combine signals in *novel* ways to describe *novel* situations.

INTERSPECIES COMMUNICATION

It was in the latter half of the 1960s through the 1970s – somewhat congruent with the period described above – that parallel sets of experiments involving both human language and nonhumans were begun. Here, researchers' goals were not to crack the code that nonhumans were using among themselves, but instead were to try to teach a variety of species, through multiple modalities, to communicate directly with humans, with the idea that such interspecies communication would be "a possible window on the minds of animals" (Griffin, 1976, ch. 7). The intent was to demonstrate that, given such training, the various species would develop true symbolic reference and at least some level of rule-governed performance (a basic form of syntax); the underlying premise was that such behavior could not be instilled *de novo*, but that it could be taught only if it were based on some already existent abilities (or even predispositions). Specifically, the extent of success in instilling symbolic reference would provide some evidence for, at the very least, some cognitive underpinnings of referential nonhuman communication systems: even if humans could not find ways to unequivocally demonstrate aspects of human language in nonhumans' native communication systems, absence of evidence might not necessarily prove evidence of absence, and with training, maybe latent abilities could be brought to light. Thus, by using interspecies communication, humans would begin to explore the cognitive and linguistic capacities of nonhumans.

Not surprisingly, such studies began with the great apes – creatures with a close phylogenetic relationship to humans (reviewed in Marks, 2005), and cetaceans – creatures with large brains relative to their body sizes (reviewed in Ridgeway, 1990).

First came several failed attempts to train nonhumans to speak English (e.g., Lilly, 1967; Kellogg, 1968; Hayes and Nissen, 1971). Later, more successful experiments followed: The ape studies involved chimpanzees (*Pan troglodytes*) trained with American Sign Language (ASL; Gardner and Gardner, 1969), magnetized plastic chips (Premack, 1971), and a computer-based system (Rumbaugh, 1977). A gorilla (*Gorilla gorilla*, Patterson, 1978) and an orangutan (*Pongo pygmaeus*, Miles, 1978) were also trained with ASL and Signed English. The early dolphin (*Tursiops truncatus*) studies involved vocal whistle imitation and responses to arm signals (Herman, 1980).

Intrigued by all these studies and armed with knowledge of the birdsong research plus the clear mimetic ability of parrots, I decided, in 1977, to determine whether a Grey parrot could also engage in a related form of interspecies communication (Pepperberg, 1999) – one using the sounds of English speech. The odds were not in my favor: Parrots, as noted above, are separated from primates by over 300 million years of evolution (Hedges et al., 1996); they were thought to be mindless mimics (Lenneberg, 1967); at the time were presumed to lack anything like a cortical area (the earliest confirmation of such a functionally homologous area was not published until Reiner et al., 2004; Jarvis et al., 2005, although some glimmerings did already exist: see Cobb, 1960; Portman and Stinglin, 1961; Nauta and Karten, 1971); and previous studies using the standard conditioning techniques of the time had failed to instill communicative competence in a variety of mimetic avian species (e.g., Mowrer, 1950, 1952, 1954; Grosslight and Zaynor, 1967). A subsequent study (Todt, 1975), however, recognized the importance of social interaction in training techniques (specifically, use of a modeling procedure called the model/rival or M/R technique, in which two humans demonstrated the types of interactive vocal behavior that the subject is to learn) and eschewed conditioning. Todt (1975) demonstrated some level of success in that his Grey parrot subject engaged in a limited number of dialogues with its human trainer. Such findings suggested that the psychologists' previous failures to achieve meaningful communication with their birds (or to find any level of complex cognitive processing) might be a consequence of inappropriate training techniques, rather than any inherent lack of ability in their psittacine subjects, and that vocal human-parrot communication might be possible. Specifically, I argued that by using additional, fairly new information on social learning (e.g., Bandura, 1971; Todt, 1975) and what little was known about parrot communication at the time (e.g., Power, 1966a,b; Nottebohm, 1970; Busnel and Mebes, 1975; Wickler, 1976; for a detailed review, see Pepperberg, 1999), I could adapt this M/R technique and achieve some level of referential acquisition.

Interestingly, all attempts at interspecies communication using human-based systems succeeded to some extent. Results from the different laboratories were divergent, but complementary (Pepperberg, 2017). The studies using variants of sign language (Gardner and Gardner, 1969; Miles, 1978; Patterson, 1978) allowed the apes to exhibit flexibility and innovation; because their system was also used with humans, it allowed direct comparisons of communicative acquisition between child and ape. Alternatively, the use of an original no-fault training procedure that rewarded associations of plastic chips with physical objects and enabled sophisticated tests of both trained and untrained chimpanzees (Premack, 1971) provided less information about communication skills than the ASL-based studies but began to elucidate how acquisition of symbolic reference could affect cognitive processing. A computer-controlled system, using a chimpanzee-sized version of a Skinner box and an artificial "language" (Rumbaugh, 1977), provided information about which basic concepts could be acquired *via* associative learning and how such learning could still allow for innovation. Herman (1980) began to show that dolphins could respond to specific cues with specific actions that demonstrated referential comprehension. My parrot started to use the sounds of English speech to identify objects, materials, colors, and shapes (Pepperberg, 1981). We believed that we were gaining valuable insights into the origins of referential communication: if creatures separated by so many years of evolution and with remarkably different-looking brains could all acquire some level of symbolic reference and regular ordering of those symbols, would not that imply the existence of some common origin or convergence?

Our resulting publications triggered approbation and condemnation in equal amounts, including from each other (Pepperberg, 2017). Unlike most controversies in science, however, discussions that began in scholarly journals (e.g., Bronowski and Bellugi, 1970; Lachman and Mister-Lachman, 1974; Terrace et al., 1979) were soon abandoned. Arguments and counter-arguments were prominently portrayed in the media, culminating in a chaotic meeting at the New York Academy of Science in the Spring of 1980 (Marx, 1980; Wade, 1980; Sebeok and Rosenthal, 1981). Important issues got lost in the resulting brouhaha, specifically questions as to, for example (Pepperberg, 2017, p. 182):

> "…what were the actual hallmarks of language, what might the apes', dolphins' and parrot's abilities tell us about language evolution and cognitive processing, what stages did children go through en route to full language, how did codes such as ASL differ from spoken language and were these differences important? (Note that at one point some scientists questioned if ASL was even a real human language; a full analysis hadn't been published until Stokoe, 1978)."

The result was that most (although not all) of us lost our funding and abandoned our emphasis on the extent to which nonhumans could acquire the elements of human language; we focused instead on using whatever levels of interspecies communication that we had instilled in our subjects to examine various forms of cognitive processing that could be specifically examined *via* symbolic reference – a single feature of human language. Thus, despite our abandonment of studies of how much and how many aspects of human-based language a nonhuman could acquire, the issue of *reference* remained basic to our findings.

REFERENTIAL COMMUNICATION

Deacon (1997, p. 44), in his influential book *The Symbolic Species*, argues that symbolic reference is "the central riddle in the problem of language origins" and claims that, except for those few nonhumans trained in interspecies communication, it is what separates human and nonhuman minds. Whether or not one agrees with his overall thesis, his assertions with respect to the importance of reference cannot easily be ignored. His point is that "reference" is not present when a label (or a sign) is simply *associated* with something (e.g., as is a red button that, when hit, delivers food whereas a green one does not) but is present if the label actually stands for something in a unique manner that is independent of context (e.g., "blue" describes the color of a pansy, a berry, a certain wavelength of light, as well as the skin color of a well-known but *fictional* entity). Once an individual understands symbolic reference at this level, the information content of symbols can be manipulated independently of their physical instantiation. Thus, finding the extent to which nonhumans' communication systems – whether natural or acquired – involve reference can be central to (i.e., affect) how they process information, and thus to their cognitive abilities (Premack, 1983). This point is one that I will discuss at length below; for now, the discussion concerns how to determine reference (i.e., and levels thereof) in a nonhuman system.

Separating reference from association is not a trivial task. Initially, biologists and linguists formulated "design features" of human language (e.g., Hockett and Altmann, 1968), in which issues, such as *arbitrariness*, *interchangeability*, *displacement*, and *semanticity*, play a role in designating what constitutes signals that are referential – rather than simple associations. Arbitrariness eliminates signals that cannot be separated from the referent, such as the meowing of a cat to designate a cat, unlike "c-a-t" in English, or "c-h-a-t" in French – if you do not know the language, you do not know the meaning. Interchangeability eliminated signals that travel only from sender to receiver, or vice-versa, like a pheromone that signals sexual receptivity; one sex emits it, the other attends, and that is the extent of its use. Displacement eliminates signals that are used only to indicate something that is physically present, that is, signals that cannot be used to describe something in the past or future, or that one would wish to be present; a food grunt does not describe the antelope that was consumed yesterday. Semanticity eliminates signals that do not designate something specific; in a Piercian sense (see Pierce, 1978), smoke indicates the presence of fire, and can be seen to "represent" fire, but can also imply many things related to fire, such as a type of meat being grilled, and thus "smoke" is not considered to be the label for fire. As noted above, researchers initially claimed reference for vervet alarm calls (Seyfarth et al., 1980a,b); after examining the issues of reference, however, these calls seem to be more indexical, in that they indicate the immediate presence of something and the type of response that one must take. It would seem, however, that such level of use is completely sufficient for the vervets; the issue for humans who are trying to establish exact levels of reference (see below) is the difficulty of, for example, designing experiments to determine whether nonhuman communication involves material such as telling one another to avoid the drinking hole near the anthill today because somebody saw a leopard there this morning. The few attempted experiments to examine possibilities of that nature (e.g., Cheney and Seyfarth, 1985) have multiple alternative explanations for the resulting data: should a vervet produce an alarm call at the sight of a gazelle carcass that humans have deceptively cached in a leopard-like manner? Although the presence of the dead animal may mean that a leopard is nearby and thus that a call is appropriate, the carcass also likely means that the leopard would already have plenty of food and will not be hunting a monkey anytime soon, such that a call is not appropriate. Again, such is not to argue that the vervet calls, or, for example, those of Diana monkeys in which the severity of an alarm call can be tempered by its use in combination with another call (Candiotti et al., 2012) are lacking informational content – the existence of content has unequivocally been established. Such communication, however, would not appear to have the highest level of *symbolic* reference. But what exactly is meant by "symbolic reference"?

Deacon (1997) devotes a large part of his book to examining what separates reference – in his words, "the symbolic threshold" – from other levels of meaning, and does so much more elegantly than I can summarize here, particularly as my main objective is *not* to define reference but rather to discuss how the acquisition of symbolic, referential communication in nonhumans may affect the ways in which they process information. For the sake of readers of this paper, my interpretation of "symbolic reference" is that it involves semantic and pragmatic use of noniconic symbols – be they auditory/ vocal, manual, or lexical – to stand for (but not be limited to) items such as physical objects and their attributes, various concepts, relations among these items and concepts, actions that can be done to or with these items, and comments about these items (e.g., relating to past/future/hypothetical issues). The use of the term "level of reference" is, again, my interpretation, and follows the above order, where the simplest level involves symbols for objects and attributes, the next level involves symbols for concepts, etc. I expand upon a few of Deacon's points that bear repeating.

As noted above, distinctions must be made between associations and reference, and such distinctions are made even more difficult given that all reference begins with associations – not in the sense that reference is built up from many associations, but that the earliest stage of learning about reference begins with learning associations: repeated correlations between the presence of the object or action *x* and hearing the sound "x." Thus, early label acquisition in children is likely more involved with associations than with actual referential meaning – hence the use of holophrases (use of a single word to indicate a variety of situations) as well as over- and under-generalizations of individual words (e.g., calling all four-legged creatures "doggie"). The very *first* label acquired by some nonhumans that are trained by humans probably is simply the association of sound or other symbol with obtaining a reward (a generalized "gimme") rather than something containing reference.

Even the first few labels are still likely simply associations between some signal and item or particular actions with particular situations. In very young children, such associative learning also occurs and generally persists through about the first 10 months, when production is minimal (e.g., Fenson et al., 1994). That is, during this period, if shown two objects – one perceptually salient and one less so – in the presence of a caretaker who focuses on and labels the less interesting object, children ignore that focus and attach/associate the label they hear to what *they* find most salient (Pruden et al., 2006). As more and more labels are acquired, and more and more associations are made, around 12 months something changes, and referential acquisition begins to occur. Social cues – the actions and focus of the person doing the labeling – start to take precedence over temporal contiguity (of label and object) and perceptual salience during acquisition; the child engages in joint attention where the adult and child both focus on *the object the adult is labeling* (see Tomasello and Farrar, 1986), sharing their experience – a communicative, *referential* act – and, by 24 months, children will ignore a more attractive item that may also be present (Hollich et al., 2000). Such behavior would not be observed if learning was, as it is at 10 months, purely associative (Golinkoff and Hirsch-Pasek, 2006). Too, for children, for example, the connection between an object (a food) and its label ("cookie") does not become extinguished when use of the label does not frequently result in obtaining the referent, as any caretaker of a toddler will attest. If the connection were mere association, extinguished use would indeed be the case (again, see Deacon, 1997; and, yes, one might argue that intermittent reinforcement might strengthen the association, but generally the intervals involved in such human communication – e.g., up to days – are considerably longer than those used in intermittent reinforcement experiments). Once reference is established, use of the label expands; for example, the label is no longer confined to referring to a specific object ("my red round bouncy thing") or even a small class of objects ("that which I use to play catch"), but can be used to identify novel instances of the item or material (e.g., a golf ball and a basketball are both recognized as "ball"; "wool" can refer to a scarf, a sweater, or even yarn) – or even used in similes and metaphors ("the moon is a balloon" – with a nod to e.e. cummings). Subjects, be they nonhuman or human, can begin to use symbols to ask for labels for novel items ("What's this?"), demonstrating an understanding of how symbols relate to one another; furthermore, hierarchical categorical labels are then learned, such that a subject knows from which particular, appropriate subset of labels to respond when asked "What color?" vs. "What shape?" (Pepperberg, 1983, 1990a), and after learning a new hue label, immediately understand its *relation* to the category "color" (and likewise for shapes, materials, numbers, etc.). Capacities for comprehension and production now become equivalent (e.g., Pepperberg, 1987b, 1990a,b; Pepperberg and Gordon, 2005). Once reference is established, subjects can use symbols to answer symbolic questions about characteristics of objects that are not immediately present, and can use symbols to request absent items (Pepperberg, 1988a, 1999). So far, although several studies in birds and apes demonstrate their ability to *plan* for the future (e.g., Kabadayi and Osvath, 2017 and references therein), no evidence yet exists for any capacity to use symbols to *refer* to the future. Thus, differences clearly exist in human vs. nonhuman levels of symbolic reference. Nevertheless, unlike nonhuman communication in the wild, for which humans have, as yet, been unable to unequivocally establish symbolic reference, nonhumans who have been taught human systems have demonstrated such reference. For me, the critical issue is one raised by Premack (1983): the extent to which such reference may affect the cognitive processes of those subjects. In the following sections, I will concentrate on my own research and leave reviews of nonhuman primate and cetacean work to researchers in those fields.

SYMBOLIC REFERENCE AND COGNITIVE PROCESSING, MOSTLY WITH RESPECT TO GREY PARROTS

Symbolic reference does not guarantee, but enables, abstract thought. Thus, an individual that can represent an object, an action, an attribute, etc., by a symbol can mentally manipulate that symbol, releasing thought processes from the here-and-now [note that an example of nonsymbolic reference would be the approximate number system (ANS) that provides a sense of quantity; the ANS allows distinguishing, e.g., between "more" vs. "less," but does not enable representation of exact quantity and thus does not enable actions such as multiplication or division). As noted above, simile and metaphor are possible; actions can be planned, tested, and altered without being physically embodied. Premack (1983) thus argued that nonhumans who learned symbolic reference have an enhanced ability to perform tasks that require abstract thinking. He buttressed these claims with data demonstrating that those of his apes that had acquired such symbolic reference outperformed those that did not. And, it was not only apes that could acquire symbolic reference – as noted above, my Grey parrots not only labeled objects, materials, attributes and requested actions, but one parrot, Alex, also used his labels to request new labels and used sound play to devise new labels (Pepperberg, 1990b; additional data reviewed in Pepperberg, 1999). He understood concepts of relative size, number, and of category (i.e., had categorical labels of "color," "shape," and "material" and understood what labels were appropriately subsumed under each; Pepperberg, 1999). Other Grey parrots, particularly one named Griffin, have also acquired symbolic reference and succeeded on various cognitive tasks (see below; e.g., Pepperberg and Nakayama, 2016; Clements et al., 2018), often outperforming subjects lacking symbolic reference. I have previously discussed several experiments from my laboratory that give additional credence to Premack (e.g., reviewed in Pepperberg, 2020a,b, 2021, in press). I summarize the importance of symbolic reference for studying cognitive processes and then briefly review some of these studies here.

In some instances, symbolic reference allows the subject to demonstrate cognitive abilities more easily (e.g., may enable them to acquire certain concepts because it allows them to think abstractly; see below) or simply makes it less difficult for humans to interpret the data. In either case, parrots' vocal plasticity allows us to evaluate their abilities because they can be tested *via symbolic interspecies communication* (Pepperberg, 1981). Interspecies communication (a) directly states the precise content of questions to be asked – animals need not determine the nature of a question through hundreds (if not thousands) of instances of trial-and-error learning, thus making the task efficient; (b) incorporates research showing that social animals may respond more readily and accurately within an ecologically valid social context (Menzel and Juno, 1985); (c) allows facile data comparisons among species, including humans; (d) is an open, arbitrary, creative code with enormous signal variety, enabling an animal to respond in novel, possibly innovative ways that demonstrate greater competence than required responses of operant paradigms, and allows researchers to examine the exact nature *and* extent of information an animal perceives; (e) allows rigorous testing that avoids expectation cuing: Subjects can be made to choose responses from their *entire* repertoire rather than from a subset relevant only to a particular topic. Interspecies communication *via* symbolic reference may thus more facilely demonstrate nonhumans' inherent capacities or enables their learning of more complex tasks. I now describe a few of several instances in which symbolic reference has been crucial in determining the extent of cognitive abilities in my Grey parrots. Additional studies have been performed for which symbolic reference has allowed testing and demonstration of competence at a level that would not otherwise have been possible (e.g., Piagetian probabilistic reasoning; Clements et al., 2018; reviewed in Pepperberg, in press), likely because such studies involve the use of symbols as abstract place-markers to assist in tasks requiring memory (note Pailian et al., 2020).

Concepts of Same-Different

A review of this entire topic is the basis for a separate paper (Pepperberg, 2021), but the central issue is as follows (Pepperberg, 1987a): Same-different is more than identity vs. non-identity or the difference in entropy – that is, in overall randomness – between stimuli sets (e.g., Young and Wasserman, 2001). Rather, it is a task that, according to the stringent criteria of Premack (1983), requires a feature analysis of the objects being compared, recognition that objects can simultaneously exhibit attributes that involve *both* similarity and difference, and the ability to understand which attributes are being targeted based on questions of either similarity *or* difference. Because an appropriate response requires that a subject (a) attend to multiple aspects of two different objects; (b) determine, from a verbal question, whether the response is to be based on sameness *or* difference; (c) determine, from the exemplars, *exactly* what is same or different (i.e., what are their colors/shapes/materials?); and then (d) produce, verbally, the label for the hierarchical category of the appropriate attribute, the task is a clear instance in which symbolic reference is likely critical for success – and one that is failed by subjects lacking such abilities (Premack, 1983). Alex succeeded in this task: he could view any two objects, even if he could not label any of their specific attributes, and produce the labels "color," "shape" or "mah-mah" (his label for matter) in response to questions of "What's same?" or "What's different?"; notably, unlike other subjects, he was not limited merely to choosing between symbols representing *same* or *different* or choosing physically between only two objects that were similar to or different from a sample (Premack, 1983) but had to produce the hierarchical category labels from a repertoire of ~70 labels. He eventually learned to respond "none" appropriately to queries about sets that were identical or completely different but only after succeeding on the initial task (Pepperberg, 1988b). By learning symbols – "same"-"different" – to represent the *relations* of categorical commonality – or lack thereof – for specific object pairs, Alex, when experiencing a novel instantiation, could likewise understand *its* relationship to the abstract representation of same-different relationships – as when, queried for the first time "What color bigger?" for two equally-sized items, he asked "What's same?" (see below, Pepperberg and Brezinsky, 1991). Such fluid response ability requires symbolic, referential, and interspecies communication.

Relational Concepts: Bigger/Smaller

Understanding relative concepts (darker than, bigger than, etc.) is a more complex task than learning to respond to an absolute concept (e.g., redness; see discussions in Schusterman and Krieger, 1986; Pepperberg and Brezinsky, 1991); it requires a subject to *compare* stimulus choices and then derive and use an underlying, more abstract (and thus general) concept. For example, learning an absolute stimulus value requires a subject to form only a single association (e.g., choose gray; Thomas, 1980), whereas in a task such as "lighter than," the subject must recognize that what is correct in one trial ("gray" in a task pitting black against gray) may be the incorrect in the next (pitting white against gray). In many tasks, subjects can acquire both absolute and relative knowledge, and because absolute knowledge is acquired more easily, the challenge to an experimenter is to demonstrate whether relative knowledge has also been acquired. Even more difficult is the simultaneous demonstration of both dimensions of relational knowledge – e.g., lighter *and* darker, bigger *and* smaller, same *and* different. A subject that uses symbolic reference, however, can simultaneously be taught labels for both concepts being tested (note Rattermann and Gentner, 1998), rather than having to derive one concept over large numbers of trials (i.e., by being rewarded for choosing only the larger) and then the other through large numbers of reversals (i.e., now being rewarded for choosing only the smaller – in this paradigm, however, *both* concepts may actually *never* be acquired, in that a subject without symbolic reference may simply learn "choose X" vs. "avoid X"; see Hochmann et al., 2016, 2018 for a discussion). Alex, after learning to respond to "What color bigger/smaller?" for three sets of items, transferred, without additional training, to a large number of sets involving sizes outside the training paradigm and to totally novel objects with respect to shape, color, and material; he also spontaneously transferred to "What matter bigger/smaller?" and, when the two objects were equal in size, spontaneously responded "none," transferring his understanding of that label from the aforementioned study on a lack of same/difference (Pepperberg, 1987a, 1988b; Pepperberg and Brezinsky, 1991). He not only responded to the largest *or* the smallest item present but also recognized that on *any* trial, *either* bigger or smaller could be queried. Such abilities are thus most clearly tested through interspecies communication systems.

Number Concepts

Almost every living creature that has been studied has demonstrated some sense of number – exact quantification for sets up to 3; approximate quantification for larger sets, for example, "more" vs. "less." In nonhumans, such abilities have been shown in creatures from fish (Petrazzini et al., 2015) to bears (Vonk and Beran, 2012); in humans, such abilities are found even in preverbal children (Wynn, 1990) and preliterate hunter-gatherer societies (e.g., Frank et al., 2008). However, subjects that understand symbolic reference can go far beyond approximation. They know that a set of *x* elements has precisely *x*, not "about *x*, ±1 or 2." That is, they can learn that individual symbols represent exact, specific quantities, whatever the items involved: a group of six ants or six elephants or six grapes have different sizes, shapes, masses, etc., but have the same number of elements. Such abilities were once thought to be limited only to humans (reviewed in Pepperberg and Carey, 2012), but a very few nonhumans have demonstrated such exact symbolic number representation, at least for quantities ≤9: two apes, Matsuzawa's Ai (Matsuzawa, 1985) and Boysen's Sheba (Boysen and Berntson, 1989), and my subject, the Grey parrot, Alex (Pepperberg, 1987b, 1994).

As I will argue in this section, symbolic reference, importantly, is a prerequisite for advanced number abilities. Although no nonhuman has, as far as we know, invented symbolic numerical representation, those that have acquired such understanding are capable of true counting and simple arithmetic capacities; they can deduce or, at the very least, *learn* cardinality and ordinality and match abilities of ~5-year-old children. I discuss Alex's abilities in detail with some references to the nonhuman primate research; a full review of the ape studies is beyond the scope of this paper and can be found in papers by Boysen, Matsuzawa, and their students.

True counting, as defined by the several counting principles ("CP," Gallistel and Gelman, 1992) is not easily acquired. CP state that numerals must be applied in order to items in a set to be enumerated and in a 1–1 correspondence, that the last numeral in a count represents a set's cardinal value, and that the successor function (that each numeral is known to be exactly one more than the one before it and exactly one less than the one after it; e.g., Carey, 2009) must be understood. This last induction allows for a "bootstrapping" process initially seen only in children. Specifically, the process by which children learn their first few numbers (1–4) is extremely slow (i.e., proceeds over the course of several years), during which time they also simultaneously learn a number line – they learn to state their numerals in a specific order – even though initially the line may make little sense and the order in which they recite their numerals can be variable (Siegel, 1982; Fuson, 1988). Eventually, the ordering of their numerals stabilizes as they learn the *symbolic* meaning of the smaller numerals and they acquire the successor function – and then the bootstrapping process engages: without any further instruction they can now immediately encode the cardinal value expressed by *any* numeral in their now stable count list; the long process used for acquiring the earlier numbers is no longer necessary. In contrast, no nonhuman had shown savings in learning as the successive numerals 5, 6, 7, etc., were added to their repertoire – that is, none had apparently induced the successor function, until Alex (see below). Interestingly, however, Alex did not learn his numerals in order (Pepperberg, 1987b), and all his labels were vocal – meaning that he had to learn not simply to point to a numeral as did the other nonhumans, but rather learn to configure his vocal tract to produce novel utterances (e.g., imagine trying to produce the /v/ sound without lips; see Patterson and Pepperberg, 1998).

Alex nevertheless acquired the ability to use his vocal English labels to quantify sets of one through six objects exactly (i.e., his accuracy did not decrease as the size of the set increased as in the case of the ANS) and was equally accurate when asked to examine novel sets and sets placed in random arrays (Pepperberg, 1987b, 1994). Such behavior is not possible without the use of symbolic reference (Pepperberg, 2020a). Furthermore, Alex, without training, was also able to quantify subsets in a heterogeneous array: given four groups of items that varied in two colors and two object categories (e.g., blue and red keys and trucks), he was able to label the number of items uniquely defined by the conjunction of one color and one object category (e.g., "How many blue key?") with an accuracy >80% (Pepperberg, 1994).

Interestingly, production and comprehension of number labels may proceed independently. Even young children who are quite proficient at producing the correct number label when asked to quantify a set often fail tests to determine how well they comprehend those labels – if given a bowl of marbles and asked to "Take four," they often grab a handful rather than the correct amount (Wynn, 1990, 1992). Unlike the other nonhuman subjects, Alex was never trained on number comprehension; nevertheless, when tested, his comprehension accuracy was somewhat superior to that of production (Pepperberg and Gordon, 2005). Again, such abilities are based on symbolic reference.

Alex also acquired other numerical capacities based on symbolic reference. After being trained to identify colored plastic Arabic numerals (1 through 6) – in the absence of any sets of items – with the appropriate vocal labels used for the corresponding numerical sets, he inferred their ordinality by responding to questions of "What color (is the) number (that is) bigger/smaller?" (Pepperberg, 2006b). He again differed from nonhuman primates, who required hundreds of training trials to demonstrate this ability. He acquired a zero-like concept; however, unlike the nonhuman primates (again), he was not trained on the concept but developed it spontaneously, using the previously acquired label "none" (Pepperberg and Gordon, 2005). Like the nonhuman primates, he also spontaneously demonstrated the ability to sum sets of small quantities and label those sets, as well as the ability to provide the label for the sum represented by the combination of Arabic numerals (i.e., recognizing the quantities these abstract symbols represented, combining those quantities, and then representing their sum as a vocal label; Pepperberg, 2006a, 2012). Moreover, after learning the labels for two additional Arabic numerals (7 and 8) in the absence of any sets of items, and their ordinal relationship to previously acquired numerals he, unlike any other nonhuman but like young children, demonstrated the ability to infer their exact cardinality (Pepperberg and Carey, 2012; for a review, see also Pepperberg, 2020b).

Optical Illusions

How do nonhumans actually *see* the world? Visual systems of most nonhumans, other than those of nonhuman primates, differ considerably from that of humans – for example, nonhumans may have much less or significantly greater color vision, or have much less or greater visual acuity, than humans; they may lack binocular overlap; their neurological architecture may be strikingly different. What exactly are the perceptual processes that are shared across species? We expect that similar evolutionary demands – visual environments, survival needs – may have led to analogous, if not necessarily homologous, solutions concerning some forms of visual processing. Parrots with a repertoire of multiple vocal responses can be rigorously tested for visual competencies, an option yet to be tried in other experimental animals (Pepperberg et al., 2008; Pepperberg and Nakayama, 2016). Specifically, the types of tasks typically used for evaluating human abilities – direct questioning about exactly what is seen – are often unsuited for research with nonverbal species, and thus direct comparisons of nonhumans with humans are not possible. Intensive training procedures were generally necessary to enable nonhumans to discriminate the initial stimulus used in visual tasks, and subjects were then tested on their recognition of similar patterns. Results thus often depended on, for example, statistical averaging over hundreds of trials of pecking/touching behavior to a very limited set of choices, and as a consequence was often highly variable and dependent upon details of the experimental design (Pepperberg et al., 2008). Nonhumans that understand symbolic reference, however, are the exception: those such as Alex and Griffin, who directly communicate with humans and can respond to the exact same stimuli as humans with the exact same responses, thus provide a unique opportunity to state exactly what they see in exactly the same way as do humans.

Given that the avian visual system is notably anatomically and neurobiologically distinct from that of humans (see review in Shimizu et al., 2010 for both similarities and differences), how might a parrot respond to common optical illusions and related visual tasks? These tasks employ early and mid-level vision, and despite neuroanatomical differences, we might expect birds and humans to respond similarly. However, data from experiments using standard operant techniques on some avian subjects were sometimes contradictory and often subject to a variety of alternative interpretations (reviewed in Regolin and Vallortigara, 1995; Pepperberg et al., 2008; Pepperberg and Nakayama, 2016). For example, subjects in these studies may have responded with respect to local cues, mass/number, or stimulus generalization (e.g., Nagasaka et al., 2007).

Interestingly, when Alex was tested on the Brentano version of the Müller-Lyer illusion and Griffin tested on amodal and modal completion (respectively, the identity of occluded and illusory Kanizsa figures), by asking them directly what they saw, they responded as did humans. The two horizontal lines in the Brentano figure were of differing colors, and Alex was asked "What color bigger/smaller?"; Griffin was shown standard colored polygons with a black circle covering one corner for amodal completion and shown black pac-men on a colored ground for illusory figures (modal completion) and in both cases asked "What shape is color-X?" (see **Figure 1**).

For the Müller-Lyer illusion, symbolic reference may merely have allowed facile testing, but for the modal and amodal tests, symbolic reference was likely a necessary factor in Griffin's ability to respond appropriately. A parrot that understands that a vocal label can represent an item, object, or action is likely to understand the representative relationship between two- and three-dimensional situations. Griffin had learned labels for shapes, and thus that a vocal label could represent an object; he could then understand how two symbols (e.g., one vocal and one visual), which separately represent the same object, can then represent each other (a form of equivalence; Pepperberg, 2006b) and thus how, for example, a three-dimensional entity can be represented by a two-dimensional drawing. That is, he could in turn deduce that the two-dimensional figures represented three-dimensional objects – one item superimposed on another (amodal) or on multiple (modal) items – so that he could appropriately identify pictures of occluded objects and Kanizsa figures (Pepperberg and Nakayama, 2016). The results of these studies (Pepperberg et al., 2008; Pepperberg and Nakayama, 2016) demonstrate how testing nonhumans that understand and appropriately use symbolic reference allows the closest possible comparison with human data, an examination of exactly how nonhumans perceive their world, and of how this perceived information is processed.

CONCLUSION

It would seem that each species has its own system for communication, each with its own unique elements and structures, that is, sufficient for its needs in nature. The human system is summarily called "language," although it, too, consists of a large number of disparate entities, each again having unique elements and structures, including those, for example, based on sign (American Sign Language, British Sign Language, etc.). One can find differences and draw parallels among the various human and nonhuman systems; doing so helps to delineate their various characteristics. The problem with doing so is that humans – despite all of our detailed work in the field for decades – are still fairly incompetent in their ability

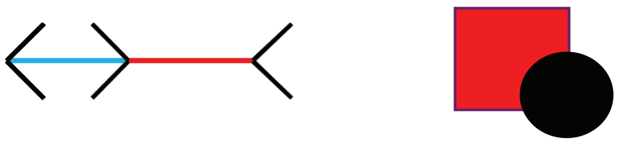


FIGURE 1 | Left to right: Brentano version of ML illusion, occluded figure, Kanizsa figure.

to perform *complete* analyses of any systems other than their own, so that additional complexities in nonhumans' systems – and thus possible additional differences and parallels with the human system – currently remain undiscovered, particularly with respect to reference (see Prat, 2019). For example, when birdsongs that are recorded at normal speed are played back at much slower speeds, many small structural differences can be observed among supposedly identical elements, emitted from different songsters or even from the same individual. Are these differences that are indistinguishable to the human ear just a bit of random noise in the system or might they carry important information to an avian listener? Humans do not yet know.

A different, although also only partially successful, tactic in examining nonhuman abilities has been to examine the extent to which nonhumans can acquire the elements of human systems – ASL, vocal labels, artificial systems built on plastic symbols or computer-based lexigrams. As noted above, the underlying premise was that such behavior could not be instilled *de novo*, but that it could be taught only if it were based on some already existent abilities (or even predispositions), such that the extent of success in instilling symbolic reference would provide some evidence for, at least, some cognitive underpinnings of referential nonhuman communication systems. At the time these studies were abruptly ended or their focus shifted (Pepperberg, 2017), no nonhuman had acquired levels of communication equivalent to those of adult humans. However, many of the nonhumans in these studies had acquired symbolic reference and, in many cases, some understanding of very simple combinatory rules for the use of these symbols. The issue of whether nonhumans understand and use such combinations – i.e., can acquire something resembling syntax – is also central for comparing human and nonhuman communication systems. However, what actually constitutes human syntax is another thorny issue, and what some researchers claim are required aspects have been shown to be lacking in some human languages (e.g., Everett, 2005). Thus, I have focused here on the symbols themselves, rather than any hierarchical organization. For a brief review of the importance of combinatorial rules in nature and those acquired by trained nonhumans, as well as their relationship to human syntax, see references cited earlier as well as additional studies and reviews such as ten Cate and Okanoya (2012), Jiang et al. (2018), and Pepperberg (in press).

The point I am trying to make is that the process of understanding that an abstract symbol can represent a concrete *item* may allow a subject to take the next step in understanding that such a symbol can also represent a *concept*, and thereby enable the subject to transfer its knowledge more easily between and among various domains. Once a subject understands that a symbol can represent a concept, the subject can mentally manipulate that symbol, releasing thought processes from the here-and-now. The subject understands how the symbols – and the concepts – are interrelated, such that they immediately understand how to use novel symbols. For example, understanding that some symbols refer to places and others to objects (i.e., representing some conceptual understanding rather than simple associations with concrete items), and that other (even somewhat similar) symbols, such as "want" vs. "wanna go" represent different classes of actions, subjects like Alex know how to use novel combinations appropriately ("I want cracker," "Wanna go shoulder") and which to avoid ("Wanna go cracker") without overt practice (note Leijnen, 2012). Subjects, such as Alex, can also apply the *concept* across domains, understanding, for example, that same-different, even though taught with respect to color/shape/material, will apply to size. Clearly, a few nonhumans, appropriately trained, have demonstrated such abilities; according to Premack (1983), these abilities are exclusively limited to those subjects receiving such training. Of course, in many instances, administering the tasks that demonstrate such abilities would be exceptionally difficult without the use of interspecies symbolic communication; thus, the extent to which such symbolic understanding is *the* critical factor enabling success is possibly still a matter for further study. However, some fairly recent studies comparing adults, young children, and nonhumans suggest that acquisition of symbolic reference – here, the labels "same" and "different" – appears to be a crucial step for being able to solve relational match-to-sample tasks (Hochmann et al., 2017). For example, some studies on *same/different* used single arrays having various mixtures of same/different icons: In some, all objects were identical; in some, all different; but for many, ratios of identical to nonidentical objects varied (e.g., Set A: 10 exemplars of one type, three of another, two of a third, and one of a fourth; Set B: four each of four different items). Unlike adult humans, who mostly responded "different" if at least one object differed from all the others – that is, by recognizing same–different relations among individual items within sets – nonhumans and 3-year-old children responded based on ratios of differing elements – on entropy, the array's overall randomness. Only when children reached about 4 years of age and began to use labels "same" and "different" appropriately did they start to respond more like adults on these types of tasks – that is, when they could rely on representations of relations among the various elements in the array; even 5-year olds were below ceiling (Hochmann et al., 2017). Thus, symbolic representation appears necessary for some forms of conceptual knowledge.

Interestingly, the comparative studies of Premack (1983) on the effects of symbolic reference solely involved chimpanzees, a species that lacks vocal learning. Might the absence of that capacity somehow be important with respect to the extent to which symbolic representation affects cognitive processing? Or might the capacity not simply for vocal learning but also for *allospecific* vocal learning be a crucial factor, because allospecific learning implies the ability to transfer concepts across systems and rapidly expand the repertoire (see Deacon, 2012)? We now know that parrots have cortical-like areas that are exceptionally large and more densely packed with neurons than those of nonhuman primates of comparable size (e.g., Jarvis et al., 2013; Olkowicz et al., 2016); that they have specific brain areas and neural connections that support extensive vocal learning – areas that appear less developed in other avian species (including parrots such as keas) that do not engage in allospecific vocal learning, and that are nonexistent in nonhuman species that lack any significant vocal learning – and that these areas also purportedly can be used to expand their intelligence (Chakraborty et al., 2015; Gutiérrez-Ibáñez et al., 2018; again, note Deacon, 2012), particularly with respect to executive function (Herculano-Houzel, 2020). Executive function involves cognitive flexibility, creative problem-solving, reasoning, and mentally representing/relating ideas and facts. Might it thus be possible that some level of reference exists in the communication systems of specific parrot species in nature? So far, evidence is lacking in other vocal learners such as songbirds for anything more than the same form of indexical reference as seen in nonhuman primates (see Beecher, this collection); in parrots, however, a possible system of individual "naming" has been discovered (Berg et al., 2012). Detailed examination of parrot repertoires is still in its infancy compared to the level of examination to which those of nonhuman primates and songbirds have been subjected; most such studies so far have done little other than describe and categorize aggressive, affiliative, and contact calls (e.g., May, 2004; Negrão de Moura et al., 2011).

Clearly, much remains to be studied about nonhuman communication systems, both in the laboratory and in the wild. Might Premack (1983) be correct about the need for symbolic reference in order to succeed on specific cognitive tasks? Or might some level of symbolic reference exist even in untrained nonhumans, providing some evolutionary communicative precursors to human language? If so, the human-nonhuman divide may not be as great as is currently imagined.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This research was supported by donors to *The Alex Foundation* and fellowships and NSF grants cited in the referenced journal articles.

REFERENCES

Adret-Hausberger, M., Güttinger, H.-R., and Merkel, F. W. (1990). Individual life history and song repertoire changes in a colony of starlings (*Sturnus vulgaris*). *Ethology* 84, 265–280. doi: 10.1111/j.1439-0310.1990.tb00802.x

Bandura, A. (ed.) (1971). "Analysis of modeling processes," in *Psychological Modeling* (Chicago: Aldine-Athcrton), 1–62.

Baptista, L. F., and Petrinovich, L. (1984). Social interaction, sensitive phases, and the song template hypothesis in the white-crowned sparrow. *Anim. Behav.* 32, 172–181. doi: 10.1016/S0003-3472(84)80335-8

Baptista, L. F., and Petrinovich, L. (1986). Song development in the white-crowned sparrow: social factors and sex differences. *Anim. Behav.* 34, 1359–1371. doi: 10.1016/S0003-3472(86)80207-X

Beecher, M. D., and Brenowitz, E. A. (2005). Functional aspects of song learning in the songbirds. *Trends Ecol. Evol.* 20, 143–149. doi: 10.1016/j.tree.2005.01.004

Beecher, M. D., Campbell, E., Burt, J. M., Hill, C. E., and Nordby, J. C. (2000). Song-type matching between neighbouring song sparrows. *Anim. Behav.* 59, 21–27. doi: 10.1006/anbe.1999.1276

Berg, K. S., Delgado, S., Cortopassi, K. A., Bessinger, S. R., and Bradbury, J. W. (2012). Vertical transmission of learned signatures in a wild parrot. *Proc. Biol. Sci.* 279, 585–591. doi: 10.1098/rspb.2011.0932

Berwick, R. C., Beckers, G. J. L., Okanoya, K., and Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Front. Evol. Neurosci.* 4:5. doi: 10.3389/fnevo.2012.00005

Bickerton, D. (1990). *Language and Species*. Chicago: University of Chicago Press.

Boysen, S. T., and Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *J. Comp. Psychol.* 103, 23–31. doi: 10.1037/0735-7036.103.1.23

Bronowski, J., and Bellugi, U. (1970). Language, name and concept. *Science* 168, 669–673. doi: 10.1126/science.168.3932.669

Busnel, R. G., and Mebes, H. D. (1975). Hearing and communication in birds: the cocktail party effect in intraspecific communication of *Agapornis roseicollis*. *Life Sci.* 17, 1567–1569. doi: 10.1016/0024-3205(75)90178-2

Byers, B. (1995). Song types, repertoires and song variability in a population of chestnut-sided warblers. *Condor* 97, 390–401. doi: 10.2307/1369025

Candiotti, A., Zuberbühler, K., and Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. *Anim. Cogn.* 15, 327–339. doi: 10.1007/s10071-011-0456-8

Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.

Catchpole, C. K. (1983). Variation in the song of the great reed warbler *Acrocephalus arundinaceus* in relation to mate attraction and territorial defence. *Anim. Behav.* 3, 1217–1225.

Chakraborty, M., Walløe, S., Nedergaard, S., Fridel, E. E., Dabelsteen, T., Pakkenberg, B., et al. (2015). Core and shell song systems unique to the parrot brain. *PLoS One* 10:e0118496. doi: 10.1371/journal.pone.0118496

Cheney, D. L., and Seyfarth, R. M. (1985). Social and nonsocial knowledge in vervet monkeys. *Philos. Trans. R. Soc. B* 308, 187–201. doi: 10.1098/rstb.1985.0019

Clements, K., Gray, S. L., Gross, B., and Pepperberg, I. M. (2018). Initial evidence for probabilistic learning by a Grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 132, 166–177. doi: 10.1037/com0000106

Cobb, S. (1960). "Observations on the comparative anatomy of the avian brain," in *Perspectives in Biology and Medicine*. *Vol. 3.* eds. D. I. Ingle and S. O. Waife (Chicago: University of Chicago Press), 383–408.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., and Townsend, S. W. (2014). Language evolution: syntax before phonology? *Proc. R. Soc. B* 281:20140263. doi: 10.1098/rspb.2014.0263

Colquitt, B. M., Merullo, D. P., Konopka, G., Roberts, T. F., and Brainard, M. S. (2021). Cellular transcriptomics reveals evolutionary identities of songbird vocal circuits. *Science* 371:695. doi: 10.1126/science.abd9704

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex. Vol. 2.* London: John Murray.

de Boysson-Bardies, B., Halle, P., Sagart, L., and Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *J. Child Lang.* 16, 1–17. doi: 10.1017/S0305000900013404

Deacon, T.W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton & Company, Inc.

Deacon, T. W. (2012). "Beyond the symbolic species," in *The Symbolic Species Evolved*. eds. T. Schilhab, F. Stjernfelt and T. Deacon (Heidelberg, London, New York: Springer), 9–38.

Dilger, W. C. (1956). Hostile behavior and reproductive isolating mechanisms in the avian genera *Catharus* and *Hylocichla*. *Auk* 73, 313–353. doi: 10.2307/4082003

Doupe, A. J., and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Ann. Rev. Neurosci.* 22, 567–631. doi: 10.1146/annurev.neuro.22.1.567

Everett, D. L. (2005). Cultural constraints on grammar and cognition in the Pirahã: another look at the design features of human language. *Curr. Anthropol.* 46, 621–646. doi: 10.1086/431525

Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D., and Pethick, S. (1994). Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59, 1–173.

Fischer, J. (2011). "Where is the information in animal communication?" in *Animal Thinking: Contemporary Issues in Comparative Cognition*. eds. R. Menzel and J. Fischer (Cambridge, MA: MIT Press), 151–161.

Frank, M. C., Everett, D. L., Fedorenko, E., and Gibson, E. (2008). Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* 108, 819–824. doi: 10.1016/j.cognition.2008.04.007

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., and Rigler, M. (1974). Development of language in genie: case of language acquisition beyond the 'critical period'. *Brain Lang.* 1, 81–107. doi: 10.1016/0093-934X(74)90027-3

Fuson, K. (1988). *Children's Counting and Concepts of Number.* New York: Springer-Verlag.

Gallistel, C. R., and Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition* 44, 43–74. doi: 10.1016/0010-0277(92)90050-R

Gardner, R. A., and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 187, 644–672. doi: 10.1126/science.165.3894.664

Golinkoff, R. M., and Hirsch-Pasek, K. (2006). Baby wordsmith. *Curr. Dir. Psychol. Sci.* 15, 30–33. doi: 10.1111/j.0963-7214.2006.00401.x

Griffin, D. R. (1976). *The Question of Animal Awareness.* New York: Rockefeller University Press.

Grimshaw, G. M., Adelstein, A., Bryden, M. P., and MacKinnon, G. E. (1998). First-language acquisition in adolescence: evidence for a critical period for verbal language development. *Brain Lang.* 63, 237–255. doi: 10.1006/brln.1997.1943

Grosslight, J. H., and Zaynor, W. C. (1967). "Vocal behavior of the mynah bird," in *Research in Verbal Behavior and Some Neurophysiological Implications*. eds. K. Salzinger and S. Salzinger (New York: Academic Press), 5–9.

Gutiérrez-Ibáñez, C., Iwaniuk, A. N., and Wylie, D. R. (2018). Parrots have evolved a primate-like telencephalic-midbrain-cerebellar circuit. *Sci. Rep.* 8:9960. doi: 10.1038/ s41598-018-28301-4

Hailman, J., and Elowson, M. (1984). *Personal Communication*.

Hartshorne, J. K., Tennebaum, J. B., and Pinker, S. (2018). A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition* 177, 263–277. doi: 10.1016/j.cognition.2018.04.007

Hayes, K. J., and Nissen, C. H. (1971). "Higher mental functions of a home-raised chimpanzee," in *Behavior of Nonhuman Primates. Vol. 4.* eds. A. M. Schrier and F. Stollnitz (New York, NY: Academic Press) (Original work published in 1956), 57–115.

Hedges, S. B., Parker, P. H., Sibley, C. G., and Kumar, S. (1996). Continental breakup and the ordinal diversification of birds and mammals. *Nature* 381, 226–229. doi: 10.1038/381226a0

Herculano-Houzel, S. (2020). Birds do have a brain cortex—and think. *Science* 369, 1567–1568. doi: 10.1126/science.abe0536

Herman, L. M. (ed.) (1980). "Cognitive characteristics of dolphins," in *Cetacean Behavior* (New York, NY: Wiley), 408–409.

Hochmann, J.-R., Carey, S., and Mehler, J. (2018). Infants learn a rule predicated on the relation same but fail to simultaneously learn a rule predicated on the relation different. *Cognition* 177, 49–57. doi: 10.1016/j.cognition.2018.04.005

Hochmann, J.-R., Mody, S., and Carey, S. (2016). Infants' representations of same and different in match- and non-match-to-sample tasks. *Cogn. Psychol.* 86, 87–811. doi: 10.1016/j.cogpsych.2016.01.005

Hochmann, J.-R., Tuerk, A. S., Sandborn, S., Zhu, R., Long, R., Dempster, M., et al. (2017). Children's representation of abstract relations in relational/array match-to-sample tasks. *Cogn. Psychol.* 99, 17–43. doi: 10.1016/j.cogpsych.2017.11.001

Hockett, C. F., and Altmann, S. A. (1968). "A note on design features," in *Animal Communication: Techniques of Study and Results of Research*. ed. T. Sebeok (Bloomington, IN: Indiana University Press), 61–72.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monog. Soc. Res. Child Dev.* 65, i–vi.

Howard, E. (1920). *Territory and Bird Life.* Glasgow, Scotland, UK: Williams Collins Sons & Company, Limited.

Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science* 366, 50–54. doi: 10.1126/science.aax0287

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159. doi: 10.1038/nrn1606

Jarvis, E. D., Yu, J., Rivas, M. V., Horita, H., Feenders, G., Whitney, O., et al. (2013). Global view of the functional molecular organizations of the avian cerebrum: mirror images and functional columns. *J. Comp. Neurol.* 521, 3614–3665. doi: 10.1002/cne.23404

Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., and Wang, L. (2018). Production of supraregular spatial sequences by macaque monkeys. *Curr. Biol.* 28, 1851–1859. doi: 10.1016/j.cub.2018.04.047

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* 21, 60–99. doi: 10.1016/0010-0285(89)90003-0

Kabadayi, C., and Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science* 357, 202–204. doi: 10.1126/science.aam8138

Kellogg, W. N. (1968). Communication and language in the home-raised chimpanzee. *Science* 162, 423–438. doi: 10.1126/science.162.3852.423

Konishi, M. (1964). Song variation in a population of Oregon juncos. *Condor* 66, 423–436. doi: 10.2307/1365432

Krebs, J. R. (1977). "Song and territory in the great tit *Parus major,*" in *Evolutionary Ecology*. eds. B. Stonehouse and C. Perrins (London: Macmillan), 47–62.

Krebs, J. R., Ashcroft, R., and Webber, M. I. (1978). Song repertoires and territory defense in the great tit. *Nature* 271, 539–542. doi: 10.1038/271539a0

Kroodsma, D. E. (1976). Reproductive development in a female songbird: differential stimulation by quality of male song. *Science* 192, 574–575. doi: 10.1126/science.192.4239.574

Kroodsma, D. E., Bereson, R. C., Byers, B. E., and Minear, E. (1989). Use of song types by the chestnut-sided warbler: evidence for both intra- and inter-sexual functions. *Can. J. Zool.* 67, 447–456. doi: 10.1139/z89-065

Lachman, R., and Mister-Lachman, J. (1974). Language in man, monkeys, and machines. *Science* 185, 871–873. doi: 10.1126/science.185.4154.871

Lawson, R. W., and Lanning, D. V. (1980). "Nesting and status of the maroon-fronted parrot (*Rhynchopsitta terrisi*)," in *Conservation of New World Parrots*. ed. R. F. Pasquier (Washington, DC: ICBP Tech. Publ. No. 1), 385–392.

Leijnen, S. (2012). "Emerging symbols," in *The Symbolic Species Evolved*. eds. T. Schilhab, F. Stjernfelt and T. Deacon (Heidelberg, London, New York: Springer), 253–262.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.

Levinson, S. T. (1980). "The social behavior of the white-fronted Amazon (*Amazona albifrons*)," in *Conservation of New World Parrots*. ed. R. F. Pasquier (Washington, DC: ICBP Tech. Publ. No. 1), 403–417.

Lilly, J. C. (1967). "Dolphin's vocal mimicry as a unique ability and a step toward understanding," in *Research in Verbal Behavior and Some Neurophysiological Implications*. eds. K. Salzinger and S. Salzinger (New York, NY: Academic Press), 21–27.

Lofting, H. (1920). *The Story of Dr. Doolittle*. New York: Frederick A. Stokes.

Lorenz, K. (1952). *King Solomon's Ring*. trans. M. K. Wilson (New York: Harper & Row).

Marks, J. M. (2005). *What It Means to Be 98% Chimpanzee: Apes, People, and Their Genes.* Berkeley: University of California Press.

Marler, P. (1952). Variation in the song of the chaffinch, *Fringilla coelebs*. *Ibis* 94, 458–472. doi: 10.1111/j.1474-919X.1952.tb01845.x

Marler, P. (1956). The voice of the chaffinch and its function as a language. *Ibis* 98, 231–261. doi: 10.1111/j.1474-919X.1956.tb03042.x

Marler, P. (1960). "Bird songs and mate selection," in *Animal Sounds and Communication*. ed. W. N. Tavolga (Washington, DC: A.I.B.S. Symposium Proceedings), 348–367.

Marler, P. (1961). The logical analysis of animal communication. *J. Theor. Biol.* 1, 295–317. doi: 10.1016/0022-5193(61)90032-7

Marler, P. (1970a). A comparative approach to vocal learning: song development in white-crowned sparrows. *J. Comp. Physiol. Psychol.* 71, 1–25. doi: 10.1037/h0029144

Marler, P. (1970b). Birdsong and speech development: could there be parallels? *Am. Sci.* 58, 669–673.

Marler, P. (1974). "Animal communication," in *Nonverbal Communication: Advances in the Study of Communication and Affect*. eds. L. Krames, P. Pliner and T. Alloway (New York: Plenum Press), 25–50.

Marler, P. (1977). "The structure of animal communication sounds," in *Recognition of Complex Acoustic Signals*. ed. T. H. Bullock (Berlin: Dahlem Workshop, Abakon-Verlagsgesellschaft), 17–35.

Marx, J. (1980). Ape-language controversy flares up. *Science* 207, 1330–1333. doi: 10.1126/science.7355293

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature* 315, 57–59. doi: 10.1038/315057a0

May, D. (2004). The vocal repertoire of Grey parrots (*Psittacus erithacus*) living in the Congo Basin (Central African Republic, Cameroon). doctoral dissertation. Tucson, AZ: University of Arizona.

Menzel, E. W. Jr., and Juno, C. (1985). Social foraging in marmoset monkeys and the question of intelligence. *Phil. Trans. R. Soc. London* B308, 145–158. doi: 10.1098/rstb.1985.0016

Miles, H. L. (1978). "Language acquisition in apes and children," in *Sign Language Acquisition in Man and Ape*. ed. F. C. C. Peng (Boulder, CO: Westview Press), 108–120.

Mowrer, O. H. (1950). *Learning Theory and Personality Dynamics*. New York: Ronald Press.

Mowrer, O. H. (1952). The autism theory of speech development and some clinical applications. *J. Speech Hear. Disord.* 17, 263–268. doi: 10.1044/jshd.1703.263

Mowrer, O. H. (1954). A psychologist looks at language. *Am. Psychol.* 9, 660–694. doi: 10.1037/h0062737

Nagasaka, Y., Lazareva, O. F., and Wasserman, E. A. (2007). Prior experience affects amodal completion in pigeons. *Percept. Psychophys.* 69, 596–605. doi: 10.3758/BF03193917

Nauta, W. J. H., and Karten, H. J. (1971). "A general profile of the vertebrate brain, with sidelights on the ancestry of cerebral cortex," in *The Neurosciences Second Study Program. Evolution of Brain and Behavior*. ed. F. O. Schmitt (New York: Rockefeller University Press), 7–26.

Negrão de Moura, L., da Silva, M. L., and Vielliard, J. (2011). Vocal repertoire of wild breeding orange-winged parrots *Amazona amazonica* in Amazonia. *Bioacoustics* 20, 331–339. doi: 10.1080/09524622.2011.9753655

Nice, M. M. (1943). *Studies in the Life History of the Song Sparrow, II*. New York: Linnaean Society.

Nottebohm, F. (1966). The role of sensory feedback in the development of avian vocalizations. doctoral dissertation. Berkeley, CA: University of California.

Nottebohm, F. (1970). Ontogeny of bird song. *Science* 167, 950–956. doi: 10.1126/science.167.3920.950

Nottebohm, F. (2006). The road we travelled: discovery, choreography, and significance of brain replaceable neurons. *Ann. N. Y. Acad. Sci.* 1016, 628–658. doi: 10.1196/annals.1298.027

Olkowicz, S., Kocourek, M., Lucan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., et al. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proc. Natl. Acad. Sci. U. S. A.* 113, 7255–7260. doi: 10.1073/pnas.1517131113

Oller, D. K. (1978). Infant vocalizations and the development of speech. *Allied Health Behav. Sci.* 1, 523–549.

Oller, D. K., Weiman, L. A., Doyle, W. J., and Ross, C. (1976). Infant babbling and speech. *J. Child Lang.* 3, 1–11. doi: 10.1017/S0305000900001276

Pailian, H., Carey, S., Halberta, J., and Pepperberg, I. M. (2020). Age and species comparisons of visual mental manipulation ability as evidence for its development and evolution. *Sci. Rep.* 10:7689. doi: 10.1038/s41598-020-64666-1

Patterson, F. (1978). The gestures of a gorilla: language acquisition in another pongid. *Brain Lang.* 5, 72–97. doi: 10.1016/0093-934X(78)90008-1

Patterson, D. K., and Pepperberg, I. M. (1998). A comparative study of human and Grey parrot phonation: II. Acoustic and articulatory correlates of stop consonants. *J. Acoust. Soc. Am.* 103, 2197–2213.

Pepperberg, I. M. (1981). Functional vocalizations of an African Grey parrot (*Psittacus erithacus*). *Z. Tierpsychol.* 55, 139–160.

Pepperberg, I. M. (1983). Cognition in the African Grey parrot: preliminary evidence for auditory/vocal comprehension of the class concept. *Anim. Learn. Behav.* 11, 179–185. doi: 10.3758/BF03199646

Pepperberg, I. M. (1987a). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. *Anim. Learn. Behav.* 15, 423–432. doi: 10.3758/BF03205051

Pepperberg, I. M. (1987b). Evidence for conceptual quantitative abilities in the African Grey parrot: labeling of cardinal sets. *Ethology* 75, 37–61. doi: 10.1111/j.1439-0310.1987.tb00641.x

Pepperberg, I. M. (1988a). An interactive modeling technique for acquisition of communication skills: separation of 'labeling' and 'requesting' in a psittacine subject. *Appl. Psycholinguist.* 9, 59–76. doi: 10.1017/S014271640000045X

Pepperberg, I. M. (1988b). Acquisition of the concept of absence by an African Grey parrot: learning with respect to questions of same/different. *J. Exp. Anal. Behav.* 50, 553–564. doi: 10.1901/jeab.1988.50-553

Pepperberg, I. M. (1990a). Cognition in an African Grey parrot (*Psittacus erithacus*): further evidence for comprehension of categories and labels. *J. Comp. Psychol.* 104, 41–52. doi: 10.1037/0735-7036.104.1.41

Pepperberg, I. M. (1990b). Referential mapping: attaching functional significance to the innovative utterances of an African Grey parrot (*Psittacus erithacus*). *Appl. Psycholinguist.* 11, 23–44. doi: 10.1017/S0142716400008274

Pepperberg, I. M. (1994). Evidence for numerical competence in an African Grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 108, 36–44. doi: 10.1037/0735-7036.108.1.36

Pepperberg, I. M. (1999). *The Alex Studies*. Cambridge, MA: Harvard University Press.

Pepperberg, I. M. (2006a). Grey parrot (*Psittacus erithacus*) numerical abilities: addition and further experiments on a zero-like concept. *J. Comp. Psychol.* 120, 1–11. doi: 10.1037/0735-7036.120.1.1

Pepperberg, I. M. (2006b). Ordinality and inferential abilities of a Grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 120, 205–216. doi: 10.1037/0735-7036.120.3.205

Pepperberg, I. M. (2012). Further evidence for addition and numerical competence by a Grey parrot (*Psittacus erithacus*). *Anim. Cogn.* 15, 711–717. doi: 10.1007/s10071-012-0470-5

Pepperberg, I. M. (2017). Animal language studies: what happened? Special issue. *Psychon. Bull. Rev.* 24, 181–185. doi: 10.3758/s13423-016-1101-y

Pepperberg, I. M. (2020a). Human-avian comparisons in cognitive performance. *Front. Psychol*, 11–973. doi: 10.3389/fpsyg.2020.00973

Pepperberg, I. M. (2020b). "The comparative psychology of intelligence: some thirty years later," in *Special Issue of Frontiers in Psychology: Comparative Psychology*. eds. D. Scarf and M. Colombo.

Pepperberg, I. M. (2020c). Vocal communication in nonhuman animals: view from the wings. *Anim. Behav. Cogn.* 7, 95–100. doi: 10.26451/abc.07.02.03.2020

Pepperberg, I. M. (2021). How do a pink plastic flamingo and a pink plastic elephant differ? Evidence for abstract representations of the relations of same-different in a Grey parrot. *Curr. Opin. Behav. Sci.* 37, 146–152. doi: 10.1016/j.cobeha.2020.12.010

Pepperberg, I. M. (in press). "Symbolic communication in the Grey parrot," in *Cambridge Handbook of Animal Cognition*. eds. A. Kaufman and J. Call (Cambridge, UK: Cambridge University Press).

Pepperberg, I. M., and Brezinsky, M. V. (1991). Relational learning by an African Grey parrot (*Psittacus erithacus*): discriminations based on relative size. *J. Comp. Psychol.* 105, 286–294. doi: 10.1037/0735-7036.105.3.286

Pepperberg, I. M., and Carey, S. (2012). Grey parrot number acquisition: the inference of cardinal value from ordinal position on the numeral list. *Cognition* 125, 219–232. doi: 10.1016/j.cognition.2012.07.003

Pepperberg, I. M., and Gordon, J. D. (2005). Numerical comprehension by a Grey parrot (*Psittacus erithacus*), including a zero-like concept. *J. Comp. Psychol.* 119, 197–209. doi: 10.1037/0735-7036.119.2.197

Pepperberg, I. M., and Nakayama, K. (2016). Robust representation of shape by a Grey parrot (*Psittacus erithacus*). *Cognition* 153, 146–160. doi: 10.1016/j.cognition.2016.04.014

Pepperberg, I. M., Vicinay, J., and Cavanagh, P. (2008). The Müller-Lyer illusion is processed by a Grey parrot (*Psittacus erithacus*). *Perception* 37, 765–781. doi: 10.1068/p5898

Peters, S., and Nowicki, S. (2017). Overproduction and attrition: the fates of songs memorized during song learning in songbirds. *Anim. Behav.* 124, 255–261. doi: 10.1016/j.anbehav.2016.09.019

Petrazzini, M. E. M., Agrillo, C., Izard, V., and Bisazza, A. (2015). Relative versus absolute numerical representation: can guppies represent “fourness”? *Anim. Cogn.* 18, 1007–1017. doi: 10.1007/s10071-015-0868-y

Pierce, C. S. (1978). *Collected Papers*. eds. C. Hartshorne and P. Weiss (Cambridge, MA: Belknap).

Portmann, A., and Stingelin, W. (1961). “The central nervous system” in *Biology and Comparative Physiology of Birds. Vol. 2*. ed. A. J. Marshall (New York: Academic Press), 1–36.

Power, D. M. (1966a). Agnostic behavior and vocalizations of orange-chinned parakeets in captivity. *Condor* 68, 562–558. doi: 10.2307/1366264

Power, D. M. (1966b). Antiphonal duetting and evidence for auditory reaction time in the orange chinned parakeet. *Auk* 83, 314–319. doi: 10.2307/4083033

Prat, Y. (2019). Animals have no language, and humans are animals too. *Perspect. Psychol. Sci.* 14, 885–893. doi: 10.1177/1745691619858402

Premack, D. (1971). “On the assessment of language-competence in the chimpanzee,” in *Behavior of Nonhuman Primates. Vol. 4.* eds. A. M. Schrier and F. Stollnitz (New York, NY: Academic Press), 186–228.

Premack, D. (1983). The codes of man and beast. *Behav. Brain Sci.* 6, 125–167. doi: 10.1017/S0140525X00015077

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2014). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5:13220. doi: 10.1038/srep13220

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., and Hennon, E. A. (2006). The birth of words: ten-month-olds learn words through perceptual salience. *Child Dev.* 77, 266–280. doi: 10.1111/j.1467-8624.2006.00869.x

Rasmussen, K. (1972). “The Netsilik Eskimos,” in *Shaking the Pumpkin*. ed. J. Rothenberg (Garden City, NY: Doubleday), 45.

Rattermann, M. J., and Gentner, D. (1998). “The use of relational labels improves young children’s performance in a mapping task,” in *Advances in Analogy Research: Integration of Theory and Data From the Cognitive, Computational, and Neural Sciences*. eds. K. Holyoak, D. Gentner and B. Kokinov (Sophia: New Bulgarian University Press), 274–282.

Regolin, L., and Vallortigara, G. (1995). Perception of partly occluded objects in young chicks. *Percept. Psychophys.* 57, 971–976. doi: 10.3758/BF03205456

Reiner, A., Perkel, D. J., Bruce, L. L., Butler, A., Csillag, A., Kuenzel, W., et al. (2004). Revised nomenclature for avian telencephalon and some related brainstem nuclei. *J. Comp. Neurol.* 473, 377–414. doi: 10.1002/cne.20118

Rice, J. O., and Thompson, W. L. (1968). Song development in the indigo bunting. *Anim. Behav.* 16, 462–469. doi: 10.1016/0003-3472(68)90041-9

Ridgeway, S. H. (1990). “The central nervous system of the bottlenosed dolphin,” in *The Bottlenosed Dolphin*. eds. S. Leatherwood and R. R. Reeves (San Diego: Academic Press), 69–97.

Rumbaugh, D. M. (1977). *Language Learning by a Chimpanzee*. New York, NY: Academic Press.

Saunders, A. A. (1951). The song of the song sparrow. *Wils. Bull.* 63, 99–109.

Schusterman, R. J., and Krieger, K. (1986). Artificial language comprehension and size transposition by a California sea lion (*Zalophus californianus*). *J. Comp. Psychol.* 100, 348–355. doi: 10.1037/0735-7036.100.4.348

Sebeok, T. A., and Rosenthal, R. (eds.) (1981). *The Clever Hans Phenomenon: Communication With Horses, Whales, Apes, and People* (New York: New York Academy of Sciences Annals).

Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Züberbühler, K., and Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Anim. Behav.* 80, 3–8. doi: 10.1016/j.anbehav.2010.04.012

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980a). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980b). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28, 1070–1094. doi: 10.1016/S0003-3472(80)80097-2

Shimizu, T., Patton, T. B., and Husband, S. A. (2010). Avian visual behavior: organization of the telencephalon. *Brain Behav. Evol.* 75, 204–217. doi: 10.1159/000314283

Siegel, L. S. (1982). “The development of quantity concepts: perceptual and linguistic factors,” in *Children’s Logical and Mathematical Cognition*. ed. C. J. Brainerd (Berlin: Springer-Verlag), 123–155.

Smith, W. J. (1963). Vocal communication of information in birds. *Am. Nat.* 97, 117–125. doi: 10.1086/282261

Smith, W. J. (1996). Information about behavior provided by Louisiana waterthrush, *Seurus motacilla* (Parulinae), songs. *Anim. Behav.* 51, 785–799. doi: 10.1006/anbe.1996.0083

Struhsaker, T. T. (1967). “Auditory communication among vervet monkeys (*Cercopithecus aethiops*),” in *Social Communication Among Primates*. ed. S. A. Altmann (Chicago: University of Chicago Press), 281–324.

Suzuki, T. N., Griesser, M., and Wheatcroft, D. (2019). Syntactic rules in avian vocal sequences as a window into the evolution of compositionality. *Anim. Behav.* 151, 267–274. doi: 10.1016/j.anbehav.2019.01.009

ten Cate, C., and Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1984–1994. doi: 10.1098/rstb.2012.0055

Terrace, H. S., Petitto, L. A., Sanders, R. J., and Bever, T. G. (1979). Can an ape create a sentence? *Science* 206, 891–902. doi: 10.1126/science.504995

Thomas, R. K. (1980). Evolution of intelligence: an approach to its assessment. *Brain Behav. Evol.* 17, 454–472. doi: 10.1159/000121814

Thorpe, W. H. (1958). The learning of song patterns by birds, with special reference to the song of the chaffinch, *Fringilla coelebs*. *Ibis* 100, 535–570.

Thorpe, W. H. (1974). *Animal and Human Nature*. New York: Doubleday.

Todt, D. (1975). Social learning of vocal patterns and modes of their applications in Grey parrots. *Z. Tierpsychol.* 39, 178–188. doi: 10.1111/j.1439-0310.1975.tb00907.x

Tomasello, M., and Farrar, M. J. (1986). Joint attention and early language. *Child Dev.* 57, 1454–1463. doi: 10.2307/1130423

Townsend, S. W., and Manser, M. B. (2013). Functionally referential communication in mammals: The past, present and the future. *Ethology* 119, 1–11. doi: 10.1111/eth.12015

Vonk, J., and Beran, M. J. (2012). Bears ‘count’ too: quantity estimation and comparison in black bears (*Ursus americanus*). *Anim. Behav.* 84, 231–238. doi: 10.1016/j.anbehav.2012.05.001

Wade, N. (1980). Does man alone have language? Apes reply in riddles and a horse says neigh. *Science* 208, 1349–1351. doi: 10.1126/science.7375943

Weeden, J. S., and Falls, J. B. (1959). Differential response of male ovenbirds to recorded songs of neighboring and more distant individuals. *Auk* 76, 343–351.

Weisman, R., and Ratcliffe, L. (1987). How birds identify species information in song: a pattern recognition approach. *Learn. Motiv.* 18, 80–98. doi: 10.1016/0023-9690(87)90024-5

Wheeler, B. C., and Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evol. Anthropol.* 21, 195–205. doi: 10.1002/evan.21319

Wickler, W. (1976). The ethological analysis of attachment. Sociometric, motivational and sociophysiological aspects. *Z. Tierpsychol.* 42, 12–28.

Wynn, K. (1990). Children’s understanding of counting. *Cognition* 36, 155–193. doi: 10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children’s acquisition of the number words and the counting system. *Cogn. Psychol.* 24, 220–251. doi: 10.1016/0010-0285(92)90008-P

Yasukawa, K. (1981). Song repertoires in redwinged blackbird (*Agelaius phoeniceus*): A test of the beau Geste hypothesis. *Anim. Behav.* 24, 114–125. doi: 10.1016/S0003-3472(81)80158-3

Young, M. E., and Wasserman, E. A. (2001). Entropy and variability discrimination. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 278–293. doi: 10.1037/0278-7393.27.1.278

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 24 December 2021
doi: 10.3389/fpsyg.2021.718414

Symbolic Signal Use in Wild Chimpanzee Gestural Communication?: A Theoretical Framework

Julia Cissewski[1]* *and Lydia V. Luncz*[2]

[1]*Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany,* [2]*Technological Primates Research Group, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*

Symbolic communication is not obvious in the natural communicative repertoires of our closest living relatives, the great apes. However, great apes do show symbolic competencies in laboratory studies. This includes the understanding and the use of human-provided abstract symbols. Given this evidence for the underlying ability, the apparent failure to make use of it in the wild is puzzling. We provide a theoretical framework for identifying basic forms of symbolic signal use in chimpanzee natural communication. In line with the laboratory findings, we concentrate on the most promising domain to investigate, namely gesture, and we provide a case study in this area. We suggest that evidence for basic symbolic signal use would consist of the presence of two key characteristics of symbolic communication, namely arbitrariness and conventionalization. Arbitrariness means that the linkage between the form of the gesture and its meaning shows no obvious logical or otherwise motivated connection. Conventionalization means that the gesture is shared at the group-level and is thus socially learned, not innate. Further, we discuss the emergence and transmission of these gestures. Demonstrating this basic form of symbolic signal use would indicate that the symbolic capacities revealed by laboratory studies also find their expression in the natural gestural communication of our closest living relatives, even if only to a limited extent. This theoretical article thus aims to contribute to our understanding of the developmental origins of great ape gestures, and hence, arguably, of human symbolic communication. It also has a very practical aim in that by providing clear criteria and by pointing out potential candidates for symbolic communication, we give fieldworkers useful prerequisites for identifying and analyzing signals which may demonstrate the use of great apes' symbolic capacities in the wild.

Keywords: symbolic communication, great apes, chimpanzees, gestures, arbitrariness, conventionalization

Edited by:
Irene M. Pepperberg,
Hunter College (CUNY),
United States

Reviewed by:
Klaus Zuberbuehler,
Université de Neuchâtel,
Switzerland
Michael Charles Corballis,
The University of Auckland,
New Zealand
Erica Cartmill,
University of California,
Los Angeles, United States

****Correspondence:***
Julia Cissewski
cissewski@eva.mpg.de

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *31 May 2021*
Accepted: *15 November 2021*
Published: *24 December 2021*

Citation:
Cissewski J and Luncz LV (2021)
Symbolic Signal Use in Wild
Chimpanzee Gestural
Communication?: A Theoretical
Framework.
Front. Psychol. 12:718414.
doi: 10.3389/fpsyg.2021.718414

INTRODUCTION

Symbolic communication is still regarded as a capacity that separates humans from other animals (e.g., Deacon, 2012), thus making us the "symbolic species" (Deacon, 1997). And it is true that language, which is a highly complex, multi-level system of symbolic communication (Deacon, 1997; Webster, 2017), can be found only in humans.

One should not conclude from this, however, that symbolic signal use is absent in other animals' natural communicative repertoires. In order to detect cases of symbolic signal use in other species, we propose avoiding a language-centered approach and concentrating instead on the basic characteristic of symbolic communication: the arbitrary and conventionalized linkage between the symbol's form (e.g., sound shape) and its meaning (concept). This follows Saussure's *arbitrariness of the sign*, that is, the distinction between *le signifiant (the signifier)* and *le signifié (the signified*; de Saussure, 1916). *Arbitrariness* is also one of Hockett's design features of language (Hockett, 1960).

The form-meaning linkage of the symbol is "arbitrary," because there is no logical or otherwise motivated connection between form and meaning. Thus, the word *book* (in its spoken or written form) does not resemble the object that it denotes. Importantly, this arbitrary linkage is not genetically determined, but is transmitted socially and is thus "conventionalized" among the members of a group (e.g., Chandler, 2017; Crystal, 2019), in this case a language community. Thus, to stay with the example, different languages use different words (for instance, *book*, *livre*, and *книга*) for the same object. The presence of communicative signals with these two characteristics, namely arbitrary form-meaning linkage and conventionalization, in non-human animals' natural communication would thus provide evidence for the existence of basic symbolic signal use. In the section "Criteria for basic symbolic signal use" we describe the application of these two characteristics to chimpanzee natural gestural communication. We emphasize the very basic nature of these criteria compared to Deacon's definition of human symbolic representation.

Intriguingly, studies in laboratory settings have revealed symbolic capacities in our closest living relatives, the great apes (e.g., Patterson, 1978; Savage-Rumbaugh et al., 1986; Greenfield and Savage-Rumbaugh, 1990; Miles, 1990), as well as in marine mammals (e.g., Schusterman and Krieger, 1984; Herman, 1987), dogs (e.g., Kaminski et al., 2004), and parrots (e.g., Pepperberg and Nakayama, 2016). For a general review, see Pepperberg (2017). The bonobos, chimpanzees, gorillas, and orangutans participating in these studies were able to acquire human-provided conventionalized arbitrary signs like lexigrams and gestures from American Sign Language (ASL). They learned both to understand them and to communicate with them. This included the combination of signs to form short utterances. Moreover, they showed cognitive abilities such as categorization (the mental grouping of objects, subjects etc. according to specific properties and for specific purposes) and decontextualization (the isolation and generalization of a mental representation from the original context). It should be said that not all specialists are convinced that the published literature demonstrates that captive apes are capable of symbolic communication; for examples of this critique and for a balanced review, see Pepperberg (2017). While in our view, the evidence from the studies of great apes in captivity points to the presence of symbolic competencies, these competencies are not obvious in the apes' communicative repertoires in their natural environment.

Here, we focus on great apes' symbolic capacities. Given the laboratory findings outside the vocal domain, we will concentrate on their natural gestural communication, as we search for evidence of symbolic signal use, and specifically in chimpanzees. Note that while symbolic communication may exist in the chimpanzees' natural gestural repertoire, the number of potential candidates reported in this paper is rather small and largely confined to two contexts (playing and mating). Note further that in laboratory studies the concepts in the human-provided abstract symbols can be narrowed down considerably. This is not possible to the same extent in the concepts underlying the potentially symbolic gestures mentioned in this paper. Nevertheless, it is not obligatory for these gestures to be associated with very narrow concepts to qualify as arbitrary and conventionalized signals. The important criterion is that the users share these concepts as a result of group-specific conventionalization.[1]

We now turn to the natural gestural repertoires of the great apes and to the criteria for defining a basic form of symbolic communication.

Great Ape Gestures

Gestures are an important element of great ape communication. They can be defined as intentional movements of body parts like hands, limbs, or the head, and body postures that are directed toward another individual, are goal-directed, motorically ineffective (toward the recipient), and receive a voluntary response (Tomasello and Call, 2007). Gestures are used by great apes in the wild (MacKinnon, 1974; Goodall, 1986; Genty et al., 2009; Graham et al., 2017) and in captive settings (e.g., Tomasello et al., 1989; Pika et al., 2003, 2005; Liebal et al., 2006). Examples of gestures in chimpanzees would be PRESENT BODY PART[2] (visual modality), TOUCH (tactile modality), and STOMP (auditory modality).

Gestures in great apes are used flexibly and in accordance with the attentional state of the recipient (e.g., Liebal et al., 2004). That is, gestures of the visual modality are more likely to be employed when the recipient is attending to the sender, and tactile gestures when the recipient is not attending; the sender may visually check the attentional state of the recipient and exhibit response-waiting. The same gesture may be used in different contexts, and a single context may elicit several different gestures (e.g., Tomasello and Call, 2007).

Interestingly, great ape gestures may involve objects in the physical environment. An example for such object-associated gestures is the auditory gesture of KNUCKLE-KNOCKING found in chimpanzees of the North group of Taï National Park (Côte d'Ivoire), which consists of the knocking of knuckles on a hard surface, for instance on tree branches (Boesch, 1995). The auditory gesture of LEAF-CLIPPING that can be observed, for instance, in the

---

[1] The same is true of human language, where we have very broad and very narrow concepts. Furthermore, in human language the degree of overlap between the speaker's intention and the hearer's understanding also varies dramatically.

[2] From here on gestures are set in SMALL CAPITALS.

chimpanzees of Mahale (Tanzania) consists of taking off parts of leaves with the mouth or fingers, thereby causing a distinctive sound (Nishida, 1980). Both auditory gestures are used for sexual solicitation in the respective group.

At the ontogenetic level, Tomasello and Call (2007, 2019) divide great ape gestures into *attention getters* and *intention movements*. *Attention getters* (e.g., GROUND SLAP) draw the attention of the audience to the sender without carrying information about the specific meaning. The recipient needs to infer this meaning from the behavior accompanying the *attention getter*. *Intention movements* are truncated forms of social behaviors (e.g., ARM RAISE as a ritualization of play hitting) and therefore do not exhibit a truly arbitrary form-meaning linkage. *Intention movements* are used in the context of the underlying social behavior and their meaningfulness normally is ensured from the context. Thus, neither *attention getters* nor *intention movements* as defined by Tomasello and Call qualify as learned arbitrary gestures.

Criteria for Basic Symbolic Signal Use – and Some Potential Candidates

In our search for symbolic signal use in our closest living relatives, we propose to identify great ape gestures that fulfill the basic criteria of *arbitrariness* of form-meaning linkage and *conventionalization* among the members of a group, as outlined in the *Introduction*.

By *arbitrariness* of form-meaning linkage, we mean the absence of any logical or otherwise motivated connection between the form and the meaning of a gesture. Arbitrary linkage is thus different from iconic linkage (where the form resembles the meaning) and also from indexical linkage (where the link to the referent can be observed or inferred; this includes pointing). Note that for *arbitrariness* it is not necessary that the form of the gesture is abstract in the sense of, for example, Arabic numbers. Rather, it would be sufficient for the form not to resemble or not be otherwise connected to the meaning of the gesture. We will see later what the possibilities might be in the case of chimpanzee gestures.

Evidence for *conventionalization*,[3] that is, the sharing of form-meaning linkages among individuals, should be sought in gestures that are learned, for instance, at the group level. These group-specific gestures, that is, gestures that are shared by some or most individuals in a group but are absent in other groups of the same species, strongly suggest social transmission (e.g., Bonnie and de Waal, 2006) rather than innateness. For a contrary view, see, for instance, Byrne et al. (2017). Group-specific gestures can be found in great apes in the wild (e.g., Whiten et al., 1999, 2001) and in captivity (e.g., Pika et al., 2003, 2005; Bonnie and de Waal, 2006; Liebal et al., 2006).[4] Examples of group-specific gestures in wild chimpanzees include the above-mentioned auditory gestures of KNUCKLE-KNOCKING and LEAF-CLIPPING. Group-specific gestures are used in particular contexts and in some cases only by defined age groups or sexes. For instance, KNUCKLE-KNOCKING is found only in males in the North group of Taï National Park and used only in the mating context (Boesch, 2012a,b).

We now turn to potential candidates for basic symbolic signal use in natural great ape communication, concentrating on chimpanzee gestures. We present them here to inspire future research and to help illustrate the theoretical framework in the section "Possible pathways to basic symbolic communication." Note that our suggestions are not based on large data sets but on the observations of long-term field researchers who observed these behaviors during their targeted data collection. There are mentions in the literature, but no systematic accounts, except for LEAF-CLIPPING (Nishida, 1980). Systematic research is needed to confirm the symbolic nature of these candidates.

Potential candidates for an arbitrary and conventionalized form-meaning linkage can, in our opinion, be observed in several group-specific gestures described for three neighboring chimpanzee groups in Taï National Park (Côte d'Ivoire): the North group, the East group, and the South group. For a map and more detailed information, see the section "NEST-BUILDING: a case study." The use of gestures differs significantly from group to group (see **Table 1**). A male chimpanzee in the Taï South group may bend together a few branches or saplings when he wants to mate with a female, while in the North group, just a few kilometers away, a male would knuckle-knock for the same purpose (e.g., Boesch, 2003, 2012a). And a young chimpanzee of the East group or the South group builds a nest to invite peers to play, while in the North group holding a leaf in the mouth would be the appropriate signal (Boesch, 2012a; Luncz and Boesch, 2015). For an overview, see **Table 1**.

The microcosm of these three habituated groups comprises no more than a few square kilometers. It is characterized by ecological similarity (Luncz et al., 2012) and by genetic relatedness between groups, which is due to migrating females[5] and extra-group paternity (Schubert et al., 2011). And yet, different group-specific gestures have evolved (Boesch, 2003, 2012a,b). **Table 2** draws out the apparently arbitrary linkage between form and meaning in these gestures.

Note first that in these group-specific gestures, different forms are used to express one and the same meaning in different groups. For example, the meaning "invitation to mate" is conveyed by the form of knuckle-knocking in the North group but by bending together branches in the South group. And

---

[3] Since the term "conventionalization" is used in linguistics to describe the spreading of a vocabulary item among members of a speech community, we adopt it here in our theoretical framework.

[4] Group-specific vocalizations in the wild have been reported for chimpanzees (Crockford et al., 2004) and for orangutans (Wich et al., 2012), as well as for chimpanzees in captivity (Watson et al., 2015).

[5] Chimpanzee females are the dispersing sex, which means that with the onset of puberty females leave their native group and integrate into a new (possibly neighboring) group. The migrating females normally adjust to the cultural givens of their new group (Luncz et al., 2012; Luncz and Boesch, 2014). In this way, the behavioral repertoires of the neighboring groups are preserved over time and cultural differences can be observed for several behaviors (Luncz et al., 2012). This includes the acquisition of group-specific gestures. This could mean that the acquisition of gestures is a flexible behavior throughout the lifetime.

TABLE 1 | Group-specific gestures in the three Taï groups.

Gesture	Form	Meaning	North group	South group	East group
NEST-BUILDING	Bending together (a few) branches or saplings	Invitation to play	–	+	+
		Sexual solicitation	–	+	–
KNUCKLE-KNOCKING	Knocking knuckles on hard surface	Sexual solicitation	+	–	–
LEAF IN MOUTH	Holding a leaf in the mouth	Invitation to play	+	–	?

+ = present (observed once a week); – = absent; and ? = limited observation time (based on Luncz and Boesch, 2015).

TABLE 2 | Arbitrary relation between form and meaning in Taï group-specific gestures.

	North group	South group	East group	North group
Form	Knocking knuckles on hard surface	Bending together (a few) branches or saplings		Holding a leaf in the mouth
Meaning	Invitation to mate		Invitation to play	

second, the same form is used to express different meanings within one group or within different groups. For example, the form of bending together a few branches conveys the meaning "invitation to mate" in the South group but "invitation to play" in the East group and in the South group. Because these gestures are not species-specific but group-specific, the arbitrary form-meaning linkage cannot be genetically determined; rather it must be socially transmitted at the group level (Boesch, 1991, 2012a). A detailed case study of NEST-BUILDING is provided below. Because there are comparatively little data concerning the LEAF-IN-MOUTH gesture, we will not include it in further analysis.

Another possible candidate for conventionalized and arbitrary form-meaning linkage would be, in our view, the auditory gesture of LEAF-CLIPPING. This gesture consists of removing parts of leaves with the mouth or fingers, thereby causing a distinctive sound (as mentioned in the *Introduction*). Nishida (1980) thought it likely that the gesture emerged from the preparation of fishing rods from leaves to catch tree-living ants, a non-social behavior.

LEAF-CLIPPING can be found in several wild chimpanzee communities where it is used in different contexts (see **Table 3** for an overview). Thus, in Mahale (Tanzania) both males and females LEAF-CLIP for sexual solicitation (Nishida, 1980). This also holds for Budongo (Uganda; Hobaiter and Byrne, 2014) and Ngogo (Uganda; Watts, 2008). In Bossou (Guinea), female chimpanzees LEAF-CLIP in varied contexts (Sugiyama, 1981). In Taï (Côte d'Ivoire), males of the South group use this auditory gesture in the context of displaying; remarkably, it

TABLE 3 | LEAF-CLIPPING in different chimpanzee communities.

Community	Sender	Context
Mahale (Tanzania)	Males and females	Mating
Budongo (Uganda)	Males and females	Mating
Ngogo (Uganda)	Males and females	Mating
Bossou (Guinea)	Females	Varied contexts
Taï (Côte d'Ivoire) South group	Males	Displaying

reappeared during an alpha-male takeover, after a gap of 2 years (Kalan and Boesch, 2018).

Note that in the case of Bossou the gesture is used in varied contexts. The recipients need to discern the meaning from the accompanying behaviors. The gesture there seems to serve as a general attention getter that is conventionalized only in the sense that it is used by females exclusively. It draws the attention to the sender without in itself conveying context-specific meaning. Therefore, in this case, it cannot be considered as a potential candidate for symbolic signal use.

Contrary to that, LEAF-CLIPPING (except for Bossou) and KNUCKLE-KNOCKING do not seem to serve merely as general attention getters to direct the recipient's attention to the sender. Rather, in addition to the attention getting component that (one could argue) is inherent to all auditory gestures, LEAF-CLIPPING and KNUCKLE-KNOCKING in themselves appear to convey information about the specific context/meaning in the respective groups. This group-specific meaning (e.g., sexual solicitation) would make further context-specific signals or clues superfluous.

Thus, these gestures would go beyond the characterization of *attention getters* given by Tomasello and Call (2007, 2019), according to which *attention getters* direct the recipient's attention to the signaler; the recipient then has to discern the intended meaning from the accompanying behavior. We take this up in the section "Semantic shifts: a new perspective on the semantics of attention getters."

Furthermore, LEAF-CLIPPING and KNUCKLE-KNOCKING do not seem to be learned individually but socially. While it cannot be excluded that, e.g., KNUCKLE-KNOCKING happens to be discovered and used by an individual to draw the general attention of conspecifics to him/herself, reports by fieldworkers confirm that in the Taï North group KNUCKLE-KNOCKING is used exclusively by young adult males, only for sexual solicitation, and has been observed across generations (e.g., Boesch, 2003, 2012a; Luncz and Boesch, 2015).

The form thus seems to be linked arbitrarily to one group-specific meaning. This linkage then would be conventionalized within the community and used and understood accordingly, even without further signals.

Boesch (2003) reports that young males in the North group use KNUCKLE-KNOCKING discreetly and repeatedly to attract females, who respond by presenting sexually. There are even instances when a different female presents to the sender although he was not looking in her direction. And significantly, sexually immature females may sexually present to the sender. That is to say, the meaning of the gesture is clearly understood by itself and this is not dependent on the sexual state of the recipient nor on the visual orientation of the sender.

In summary, we have proposed several candidates for learned arbitrary form-meaning linkage that appear to exhibit the criteria of *arbitrariness* and *conventionalization*. These signals thus, in our view, could be considered as potential candidates for a basic form of symbolic communication. Systematic field research is needed to confirm this view.

In the following section, we discuss possible pathways for the emergence of basic symbolic signal use in wild chimpanzee gestural communication. We propose that conventionalized arbitrary gestures can arise ontogenetically by borrowing the form of an existing gesture or the form of a non-social behavior that acquires communicative meaning. In both cases, the resulting (group-specific) gesture is used in a different context from that of the underlying gesture or non-social behavior, thus resulting in an arbitrary form-meaning linkage.

In the section "NEST-BUILDING: a case study," we illustrate how a basic symbolic signal could emerge and operate. The visual gesture of NEST-BUILDING in chimpanzee natural gestural communication has not been described in detail in the literature so far. It can be observed in two chimpanzee groups in Taï National Park (Côte d'Ivoire). The gesture consists of bending together a few branches or saplings, and the possible contexts are mating and/or playing, depending on the group that uses it.

POSSIBLE PATHWAYS TO BASIC SYMBOLIC COMMUNICATION

We now explore how an arbitrary form-meaning linkage in the natural gestural communication of chimpanzees could come about, and how it could be conventionalized at the group level. First, we propose two routes for the emergence of arbitrary gestures within an ontogenetic time frame. We do so from a linguistic perspective.

Two Routes to Arbitrariness

Emergence of Semanticity: A New Perspective on Non-social Behaviors in Gestural Ontogeny

We suggest that learned arbitrary gestures can emerge from *non-social behaviors* that acquire communicative functions. As described in the section "NEST-BUILDING: a case study," the way that the "play nests" and the "mating nests" are built suggests that PLAY-NEST BUILDING and MATING-NEST BUILDING are based on the non-social behavior of nest building for resting. By non-social behaviors, we mean functional behaviors that are displayed outside social interactions and without a communicative purpose. That is, a non-social behavior (nest building for resting) may have developed into a social behavior and – by acquiring communicative meaning – passed to the gestural level. Note that the resulting gestures would exhibit an arbitrary form-meaning linkage, because the gestures are used in different contexts (playing/mating) from that of the underlying non-social behavior (resting). The same seems to hold for LEAF-CLIPPING, as summarized in **Table 4** and described in the section "Criteria for basic symbolic signal use – and some potential candidates."

TABLE 4 | Group-specific gestures potentially based on non-social behaviors.

Gesture	Underlying non-social behavior	Old context	New context
NEST-BUILDING	Building nests	Resting	Mating/playing
LEAF-CLIPPING	Preparing leaf mid-ribs	Foraging	Varied (see **Table 3**)

This development needs to be distinguished from *phylogenetic ritualization* (Darwin, 1872; van Hooff, 1972, 2012; Krebs and Dawkins, 1984), where the form of non-social behaviors can be "borrowed" to serve a communicative function (*principle of derived activities*, Tinbergen, 1952). Over evolutionary time, *phylogenetic ritualization* results in species-specific gestures (gestural phylogeny). An example could be the dominance signal of MOUNTING in monkeys that may have evolved from mating behavior (Liebal and Call, 2012).

In contrast to this, we propose the emergence of new gestures from non-social behaviors within a much shorter time frame (gestural ontogeny). We call this development *emergence of semanticity*; here *semanticity* denotes the meaningfulness of communicative signals, one of the universal design features of human language as identified by Hockett (1960). This means that every communicative signal consists of a form and an associated meaning (de Saussure, 1916). The resulting gestures would be shared not at the species level (as in *phylogenetic ritualization*) but at the group level.

Theories of great ape gestural ontogeny so far heavily concentrate on the ritualization of social behaviors into gestures in social interaction (for instance, *ontogenetic ritualization* as proposed by Tomasello (1996) that results in *intention movements*, see below). Non-social behaviors are under-represented in these approaches. We now turn to the mechanism which may underlie this intriguing phenomenon.

The *emergence of semanticity*, as shown in **Figure 1**, comprises, in a first step, the recombination of the form of a non-social behavior B (form$_B$: e.g., nest-building) with the meaning of a context-specific signal S (meaning$_S$: e.g., sexual solicitation/play invitation) on the sender's side. These context-specific signals carry a message about a communicative interaction. For chimpanzees, for instance, in the play context this might be a play face and/or play gait (van Hooff, 2012; see Wilson, 1975 for other species). In the mating context, the presenting of an erect penis defines the communicative context and thus determines the meaning of any other signals used in combination by the male, for instance, BRANCH SHAKING and STOMPING.[6]

---

[6] Note that BRANCH SHAKING and STOMPING under these circumstances only initially serve to get the attention of the female. Further on, after the female has noted the erect penis, these signals express the male's insistence and his motivation to trigger the female into mating with him. The term "attention getter" often used for these signals therefore is misleading, which is also pointed out by Liebal and Call (2012)

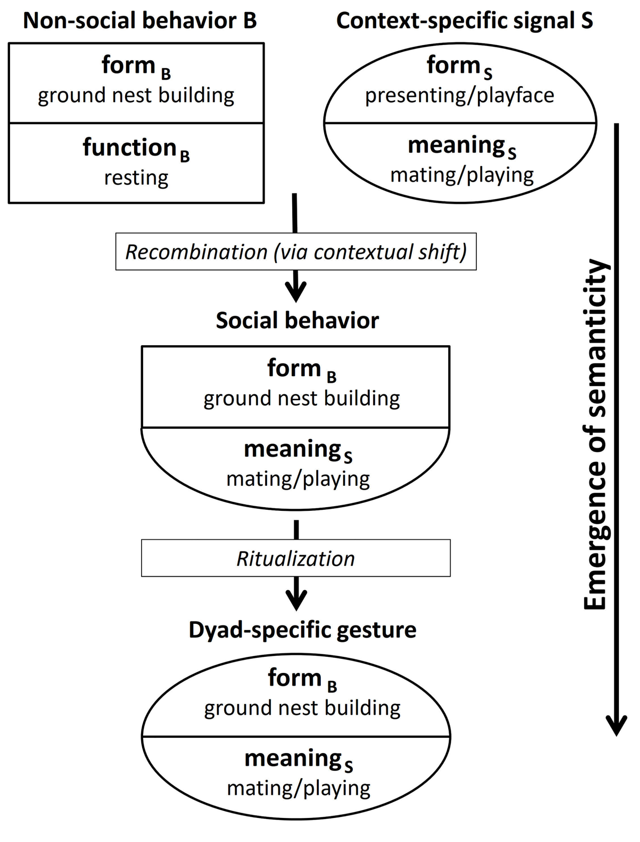


FIGURE 1 | Emergence of a dyad-specific gesture from a non-social behavior.

Thus, one could argue that the meaning of these context-stressing signals is imposed onto all signals sent at the time. However, signaling normally does not happen in a vacuum. It is not separate from other processes going on simultaneously in the environment of the sender and the recipient. Therefore, non-social behaviors exhibited by the sender that have nothing to do with the current communicative context may be drawn into the contextual field and become "colored" with context-specific meaning. Of course, it is partly a matter of chance which of the myriad of potential behaviors happening in parallel with communicative interactions (or temporarily close enough to them) are associated with the communicative context, so that the behavioral form is recombined with the meaning of the communicative signals being used.

At first, the sender may include the behavior just "because," that is, because it happened to be part of a successful interaction, even if it was not meant to be communicative. Repeated successful use of the behavior in connection with other context-specific signals may then result in a recombination of form$_B$ (provided by the nest-building behavior of the sender) and meaning$_S$ (provided by one or more context-specific signals) on the sender's side.

The new form$_B$-meaning$_S$ combination would result from a contextual shift. In the cases of PLAY-NEST BUILDING and MATING-NEST BUILDING, this means a shift from the context of resting to the context of playing/mating. In the case of LEAF-CLIPPING, the context of the underlying non-social behavior (fishing for ants) is foraging, while the resulting gesture is used, for instance, in the mating context, as illustrated in **Table 4**. The form$_B$-meaning$_S$ combination can be regarded as a social behavior (that is, socially directed by the sender toward a recipient) although in its form being based on a non-social behavior.

In a second step, this social behavior is then ritualized into a dyad-specific gesture. By dyad-specific gesture, we mean a gesture that arises within a particular dyad and is used by one or by both individuals. This ritualization is a social process that takes place within an ontogenetic time frame. During the ritualization process, the behavior may get abbreviated/truncated as proposed for *intention movements* that are based on social behaviors.

The ritualization process in the *emergence of semanticity* differs from *ontogenetic ritualization sensu* Tomasello (1996). *Ontogenetic ritualization* results in *intention movements* whose forms still represent part of the underlying (social) behaviors. Thus, the form and the meaning of gestures resulting from *ontogenetic ritualization* are connected logically and not arbitrarily. The *emergence of semanticity*, in contrast, results in gestures that do not exhibit a logical but rather an arbitrary connection between the form and the meaning of the gestures. As explained above, the reason for this phenomenon lies in the fact that the resulting gesture is used in a different context from that of the underlying non-social behavior (contextual shift).

In all examples listed above, a non-social behavior may have acquired communicative meaning. The resulting gestures would disappear with the individuals that use them, or even at some point within the individuals' lifetime. To survive, they need to be copied by other group members and thus develop into group-specific gestures. We analyze this next step in the following section "Conventionalization."

In summary, we suggest that ontogenetically arising chimpanzee gestures can be based on non-social behaviors. In this process, the form of non-social behaviors is recombined with the meaning of co-occurring context-specific signals, resulting in arbitrary form-meaning linkage. Note that in the case of NEST-BUILDING, the form of the underlying non-social behavior would have been truncated. One could thus argue that it resulted in an *intention movement* – but with arbitrary form-meaning linkage. The mechanism of *emergence of semanticity* hence would take further the concept of *intention movement* defined by Tomasello and Call (2007, 2019) by giving it an arbitrary form-meaning linkage.

Semantic Shifts: A New Perspective on the Semantics of Attention Getters

As Cissewski and Boesch (2016) have proposed, great apes may use *semantic shifts* to express new meanings without creating new forms. That is, within a community the meaning of an existing gesture would change without the form of the gesture being modified. This mechanism may underlie the group-specific usage of auditory gestures such as LEAF-CLIPPING (changing from a general attention-getter to a context-specific gesture). Importantly, the original meaning (general attention getter) of the gesture disappears, see **Figure 2**.

The reason for an existing group-specific gesture undergoing a semantic shift may be a gap in the communicative repertoire that needs to be filled, for instance, in communication under time pressure or in environments with restricted visibility. Using the gesture with the new (more specific) meaning under these circumstances would provide the sender and the recipient with adaptive benefits (for an example, see the section "Conventionalization"). A communicative gap thus can act like a vacuum that pulls existing elements into a different position in the communicative repertoire.[7]

Cissewski and Boesch (2016) argued that this phenomenon can be observed especially in auditory gestures (for instance, LEAF-CLIPPING and KNUCKLE-KNOCKING), as in this gestural modality form and meaning are less closely linked than in visual or tactile gestures that may result from *ontogenetic ritualization*. However, if a gesture of the visual or the tactile modality already exhibits arbitrary form-meaning linkage, then a semantic shift might become more feasible/likely. This could have been the case for NEST-BUILDING (see the case study below).

Excursus

The case of NEST-BUILDING is even more interesting, because in the South group the gesture is used in two different contexts (see **Table 5**), each specific to an age group. So far, we assumed that the ground-nest gestures used in the play context and in the mating context in the Taï South group emerged independently. However, it is also possible that one is based on the other. This would mean that either the ground-nest gesture used by adults in the mating context was copied by infants for the play context, or vice versa. Thus, the meaning of an existing group-specific gesture would have been modified – within an age-group. However, the original meaning is kept in the adults (or the infants, respectively), see **Figure 3** below.[8] This would mean that the semantic shift is age-group specific and that it is only partial. Both meanings exist in parallel.[9]

Note that the mechanism of group-specific semantic shifts takes further the concept of attention getters defined by Tomasello and Call (2007, 2019), by adding context-specific meaning. The gesture thus does not only draw the recipient's attention to the sender, but at the same time also includes the information as to why the attention is sought (e.g., sexual solicitation), without additional behavioral cues. The resulting gesture is not an intention movement either, but a gesture with arbitrary form-meaning linkage. Systematic field research is needed to establish that context-specific meaning is communicated by these gestures themselves, without any other context-specific signals being present (or discernable to the recipient).

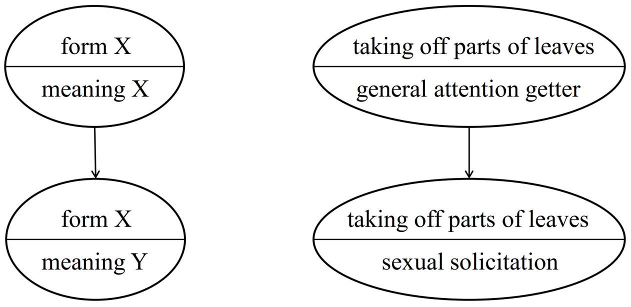


FIGURE 2 | Group-specific semantic shift.

In summary, we suggest that in some cases it is possible that attention getters undergo further development, by acquiring context-specific meaning. This would result in an arbitrary form-meaning linkage that is conventionalized first at the dyadic level and then at the group level (as discussed in the following section).

Conventionalization

The second criterion for basic symbolic signal use is *conventionalization*. True symbols cannot be innate but must be learned, in order to be shared by the members of the group. *Phylogenetic ritualization* (Darwin, 1872; van Hooff, 1972, 2012; Krebs and Dawkins, 1984), as described above in the section on the *emergence of semanticity,* results in species-specific gestures and thus would not provide an explanation for the existence of group-specific gestures. *Ontogenetic ritualization* (e.g., Tomasello, 1996), as also described in the section on the *emergence of semanticity*, results in gestures shared within dyads; it may take place in parallel in different dyads of a group, based on the same functional actions. The hypothesis of *ontogenetic ritualization* has been challenged, for instance by Genty et al. (2009) and Hobaiter and Byrne (2011). However, there is recent evidence in support of *ontogenetic ritualization* in bonobos (Halina et al., 2013). For *ontogenetic ritualization* to result in stable group-specific gestures across generations, gestures would need to be ritualized over and over again (see Genty et al., 2009). Cartmill and Hobaiter (2019) thus suggest experiments for testing whether gestures resulting from *ontogenetic ritualization* would be transferable to a *new* partner, that is, outside the original dyad. Byrne et al. (2017) propose the innateness of the majority of great ape gestures; this includes group-specific gesture as being the result of different developmental environments.

Another possibility for the rise of group-specific gestures would be social transmission. Unfortunately, little is known

---

[7] In language, whole chains of shifts can emerge in this way. Thus, some changes in the phonological inventories and lexicons of languages are commonly interpreted in this manner (e.g., the Great Vowel Shift for English).

[8] Here, we illustrate one of the two possible scenarios just mentioned.

[9] In the East group there is no evidence for the ground-nest gesture being used in the mating context. Given the lack of diachronic data, it is impossible to find out retroactively why this is the case. Either the ground-nest gesture never emerged in the mating context. Or it emerged and subsequently was lost, similar to the use of cushion-making by a limited number of male chimpanzees in the mating context in Mahale (Tanzania; Nishida, 1987; Boesch, 1995), which was used over several months, but never became fully established, and eventually vanished.

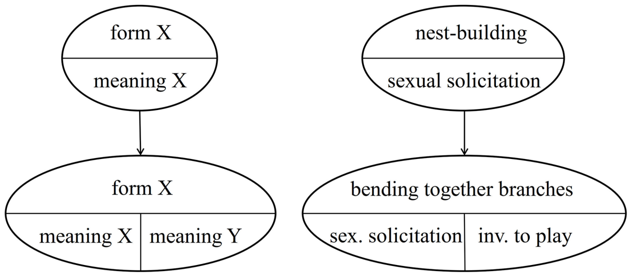


FIGURE 3 | Partial group-specific semantic shift.

about the precise mechanisms of the spreading of new gestures within groups of great apes (for a review, see Liebal et al., 2019). In view of the richness of social learning mechanisms (e.g., Hoppitt and Laland, 2008, 2013), the precise process of the conventionalization of a gesture at the group level cannot be determined retrospectively.

One could argue that the mechanism underlying the spread of NEST-BUILDING, LEAF-CLIPPING, and other group-specific gestures is observational learning, given the evidence of observational learning in captive apes (e.g., Whiten et al., 2004). Moreover, as pointed out in footnote 5, migrating females normally adjust to the cultural givens of their new group (Luncz et al., 2012; Luncz and Boesch, 2014). This includes gestures. We suggest that observational learning is a worthwhile hypothesis for field researchers to investigate further.

In practice, in the mating context, seeing the demonstrator being successful with a mating partner, should be sufficient motivation for the observer to learn the gesture. There are two variants here. The observer male might adopt the behavior specifically to attract the same female who reacted to it with the demonstrator, and this could be successful. Or the male might adopt the behavior with a different female, who has not seen it previously in this context. However, this is still a strategy which may work, since context specific-signals accompanying the new gesture would define the context. Through usage in repeated interaction, the gesture would be associated with mating in an increasing number of individuals and eventually become group-specific. In the play context, if the demonstrator successfully attracts playmates, the observer is likely to be motivated to adopt this behavior.

Note that we are not dealing with *response facilitation*, that is, NEST-BUILDING, LEAF-CLIPPING, and KNUCKLE-KNOCKING do not need the presence of a demonstrator to be displayed in the appropriate context in every-day social interaction. And we are not dealing with *program-level imitation*, because there is no novel organization of several preexisting components happening. Further note that, intriguingly, in gestures resulting from *the emergence of semanticity* or from *semantic shifts*, the form of the gesture would already have been part of the behavioral repertoire, either belonging to a non-social behavior or to a gesture. The form thus would not need to be learned.

Why then would group members adopt new gestures? As we have already mentioned, gestures may spread within a group because they provide the sender and/or recipient with adaptive benefits under specific social and ecological circumstances. Thus, to stay with the visual gesture of NEST-BUILDING, the building of mating nests enlarges the number of gestures available for sexual solicitation and thus can increase the level of persistence. Using and understanding the gesture might therefore provide an adaptive benefit in the mating context.

As proposed by Cissewski and Boesch (2016), in habitats where visibility is restricted, group-specific *semantic shifts* in auditory gestures may result in more effective communication. For instance, when signaling under time pressure, the rapid communication of meaning *via* the auditory modality provides an adaptive benefit for the signaler and/or the recipient. This could be relevant in the mating context, when mating access for males is mainly controlled by dominants while female choice is limited by male coercion. Conventionalized inconspicuous KNUCKLE-KNOCKING lets subordinate males gain mating opportunities and females gain female choice. These are strong adaptive reasons for conventionalization. Moreover, according to anecdotal evidence from Taï (Deschner, personal communication), the audience moves away from the sender when hearing LEAF-CLIPPING, because they expect an upcoming display. This reduces the risk of confrontation for the audience – and the sender.

Crucially, the effect of adaptive benefits is not strong enough to ensure that all communities with similar material and social environments converge on the same group-specific gestures. Great apes and other nonhuman primates live in complex material and social environments (e.g., Milton, 1981; Russon and Begun, 2004; Cheney and Seyfarth, 2007). The forest is no laboratory with controlled conditions. It is complex with many factors acting and interacting.

Therefore, in addition to gaining/providing adaptive benefits, we should allow for the possibility that new gestures or other behaviors may get copied without an obvious adaptive benefit, but simply because this is "how it is done." This might be the case for the generalization of PLAY-NEST BUILDING in the Taï South and East groups. Comparable scenarios have already been reported in the literature. Thus, van Leeuwen et al. (2014) report on the spontaneously emerged tradition of "grass-in-ear behavior" in one chimpanzee group of the Chimfunshi Wildlife Orphanage (Zambia). A female repeatedly put a piece of grass in her ear and left it there. Soon, other group members copied this behavior which does not have any apparent adaptive value. Another instance would be the copying of the individual-specific manner of back scratching performed by a chimpanzee with snare-damaged hands in the Ngogo community of Kibale National Park (Uganda); the copying by group members without the injury did not seem to be adaptive (Hobaiter and Byrne, 2010).

In such cases, the copying of the new behavior seems to result from a general predisposition to copy. This predisposition might have been selected for, because in itself it provides an adaptive benefit because it allows for useful behaviors to be acquired. However, the specific behaviors copied may not in every case provide an adaptive benefit. Thus, we propose that the emergence of a particular group-specific behavior, including gestures, does not need to be driven directly by adaptive benefits.

NEST-BUILDING: A CASE STUDY

We now discuss in detail one of the potential candidates for basic symbolic signal use: NEST-BUILDING. This intriguing phenomenon is found in two of the three study groups in Taï National Park (see **Figure 4**) and has not been observed in other wild chimpanzee communities. So far it has not been studied systematically nor been described in detail in the relevant literature (for mentions in the literature, see, for instance, Boesch, 2012a and Luncz and Boesch, 2015). We therefore strongly encourage fieldworkers to undertake systematic data collection and analysis to test our hypotheses.

These groups are fully habituated to the presence of humans and have been continuously observed since 1983 (Boesch and Boesch-Achermann, 2000). They engage in frequent violent intergroup encounters (Samuni et al., 2017), which do not allow them the opportunity to observe the daily behavior of members of the other groups (Boesch et al., 2008).

The gesture of NEST-BUILDING consists of bending together a few branches or tree saplings. Note that NEST-BUILDING is not to be confused with the purely functional and non-social behavior of nest-building for resting, although it is possible that the gesture borrowed its form from this non-social behavior (as proposed for the *emergence of semanticity*). Further note that the very simple constructions resulting from the gesture of NEST-BUILDING do not resemble real nests like those constructed for resting. This is clear in the video material provided for illustration. To emphasize this important distinction and to avoid misunderstandings, we first briefly describe the non-social behavior of nest building for resting and then we discuss in detail the actual gesture of NEST-BUILDING.

Nest-Building for Resting: A Non-social Behavior

Wild chimpanzees build nests, for sleeping during the night, and for resting during the day (for reviews, see Fruth and Hohmann, 1996 as well as Hicks, 2010). Day nests for resting are normally simpler than the more elaborate night nests. The chimpanzee communities of the Taï National Park follow this pattern, and day nests are commonly constructed for resting (Boesch, 1995). Although simpler than the night nests, these day nests are built by bending branches and/or saplings together, interweaving them and adding torn twigs and branches. Day nests are usually built on the ground, though sometimes also in the trees. The resting nests are normally used by one chimpanzee at a time, unless a mother has a dependent offspring. They are not used as sites for play or mating.

Video 1 illustrates the building of a day nest in a tree for resting (note that the individual briefly interrupts the building process in order to retrieve food that has been accidentally dropped).

> Video 1: https://share.eva.mpg.de/index.php/s/noTsjAJCmrRs6cm (Copyright: Liran Samuni, Taï Chimpanzee Project).

Again we emphasize that the building of day nests in the resting context is a non-social behavior. That is, it is a functional behavior that is not directed at other individuals and thus lacks communicative intent. It takes place without monitoring the attention of others, without waiting for a response and without receiving a response from other individuals.

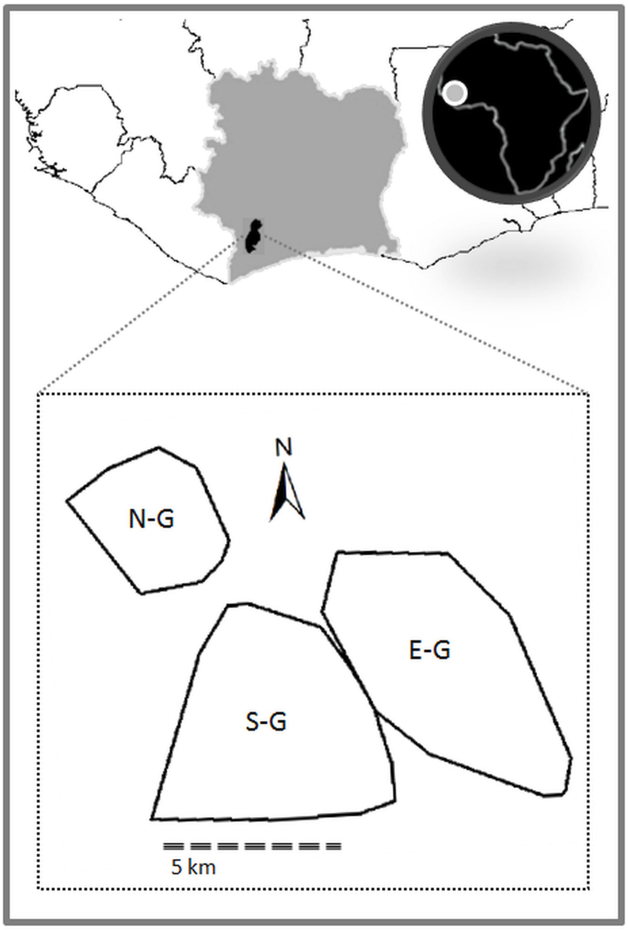


FIGURE 4 | The three habituated chimpanzee communities in Taï National Park, Côte d'Ivoire (N-G = North group; S-G = South group; and E-G = East group). Polygons indicate the home ranges of the chimpanzee groups at time of observation (2007–2009).

Nest-Building in Communicative Interaction

As we have seen, chimpanzees build nests for resting, which is a merely functional and non-social behavior. In addition, in the Taï East group and the Taï South group they may exhibit the activity of bending together vegetation with communicative intent.

The resulting gesture is called NEST-BUILDING, because the form of the gesture resembles the motorics of the act of nest-building for resting. However, the gesture does not result in a full nest but in something much simpler (see Videos 2 and 3). Note also that it is not the resulting "play nests" and "mating nests" themselves but the actual process of bending together the small number of branches/saplings

TABLE 5 | Properties and distribution of NEST-BUILDING.

Gesture	Form	Meaning	Sender	North group	South group	East group
PLAY-NEST BUILDING	Bending together a few branches or saplings	Invitation to play	Juveniles and adolescents of both sexes	−	+	+
MATING-NEST BUILDING		Sexual solicitation	Adult males	−	+	−

that has acquired communicative meaning. Similarly, the auditory gesture LEAF-CLIPPING, described above, consists precisely in the act of taking off parts of leaves and not in the bare mid-ribs that result.

In the Taï South group, NEST-BUILDING occurs in the mating context and in the play context (MATING-NEST BUILDING and PLAY-NEST BUILDING), and in the Taï East group in occurs in the play context (PLAY-NEST BUILDING; Boesch, 2012a; Luncz and Boesch, 2015). See **Table 5** for an overview. A detailed description is provided in the following.

Note that the building of nests does not serve any function in the play or in the mating contexts in the Taï South and East groups (e.g., Luncz, personal observation) outside the communicative interactions described in the following. This is important with regard to the arbitrariness of the form-meaning linkage of these gestures. Outside communicative interactions, the form (bending together a few branches or saplings) is not linked to the mating or the playing context. Thus, the nest is not linked in form to the response of the recipient or to the subsequent behavior of the signaler. The form-meaning linkage in MATING-NEST BUILDING and PLAY-NEST BUILDING would thus be truly arbitrary.

PLAY-NEST BUILDING

In order to initiate play, juvenile and adolescent chimpanzees in the East group and the South group are frequently seen bending a few surrounding saplings or branches together (e.g., Boesch, 2012a; Luncz and Boesch, 2015; e.g., Crockford et al., personal communications). Even though this behavior is observed frequently by different field researchers, there are as yet no systematic data on the use of PLAY-NEST BUILDING.

However, from September 2007 to November 2009 data were collected opportunistically by Luncz during focal follows of adult individuals, resulting in 44 independent observations of PLAY-NEST BUILDING in the East and the South groups[10]; there were, in addition, many more instances of PLAY-NEST BUILDING which were not recorded, because the researcher was focusing on adult individuals. The gesture was observed in juveniles and adolescents, both male and female, aged from 2 up to 12 years to initiate play. Both sexes responded to such play invitations. PLAY-NEST BUILDING was most frequently observed during the resting times of adult group members, a period when offspring play time is increased. Unlike resting nests, these "play nests" do normally not leave any physical evidence after play as the saplings generally regain their original structure. The saplings usually only get bent and not broken.

The sender bends together a few branches or saplings in proximity to a potential play partner (at a clear visual distance of approximately 1–5 m), taking into account the recipient's attentional state. The builder may exhibit visual checking toward the potential recipient. During or immediately after construction, which in general takes only a few seconds, the selected play partner may join the builder by interrupting him/her and play begins. Hence, the sender receives a voluntary response, that is, the potential recipient is not pulled into the nest. The building is mechanically ineffective toward the recipient. If the play partner does not react to the invitation during construction or immediately after, the builder usually sits down on the bent-over branches and looks at the potential play partner, thus exhibiting response waiting. If still nothing happens, a second round may be started or a different strategy be applied (e.g., pulling the other's leg). Note that due to the usually almost immediate reaction of the play partner, markers of intentionality like persistence or elaboration on the sender's side (e.g., by adding a second round) are hardly ever needed. Importantly, the bent-over branches clearly do not serve the purpose of resting as young chimpanzees were never observed to lie down on them. Thus, the construction resulting from PLAY-NEST BUILDING is not perceived as a nest and is not occupied by both.

Video 2 shows an example of PLAY-NEST BUILDING. Two infants play in an old resting nest in a tree. One stops and leaves the nest. The other reacts with PLAY-NEST BUILDING, and the first individual accepts the invitation and play is resumed.

Video 2: https://share.eva.mpg.de/index.php/s/tRWWbbLHAAzYcjR (Copyright: Liran Samuni, Taï Chimpanzee Project).

The reduced building process emphasizes the communicative intent, being clearly distinguishable from the original underlying behavior (nest building for resting, as shown in Video 1), especially given that it is carried out while the adult individuals are resting.

Importantly, PLAY-NEST BUILDING is interpreted as play invitation also in the absence of play-context specific signals like the play face (Luncz, personal observation; Christophe Boesch, personal communication). The gesture thus does not need any pragmatic support and serves as play invitation in its own right. PLAY-NEST BUILDING as a gesture for play invitation thus seems to be truly referential. Of course, it needs systematic data collection, ideally through video recordings, to provide firmer empirical evidence for the independent use of the gesture.

---

[10] Data collection in the Taï National Park was non-invasive and was carried out in compliance with the requirements and guidelines of the "Ministère de l'enseignement supérieure et de la recherche scientifique" and it adhered to the legal requirements of the Côte d'Ivoire. Furthermore, the regulations of the Deutsches Tierschutzgesetz (German Animal Welfare Act) and the American Society of Primatologists (ASP) principles for the ethical treatment of non-human primates were strictly adhered to.

The points listed above clearly differentiate PLAY-NEST BUILDING from the building of nests in the resting context; the latter takes place without monitoring the attention of other individuals, does not include response waiting, and is finished without receiving a response from other individuals. As detailed above, nest-building serves no function in the play context outside the communicative interactions described here. This suggests that the form of the gesture (bending a few saplings or branches together) and its meaning (play invitation) are linked arbitrarily; that is, there would be no logical connection between the two. This arbitrary linkage would be shared and thus conventionalized at the group level. Note that PLAY-NEST BUILDING has been observed for about two decades, showing its sustained use over generations and in this way providing evidence for acquisition *via* social learning. Thus, PLAY-NEST BUILDING, in our view, can be regarded as a potential candidate for investigating symbolic signal use.

MATING-NEST BUILDING

In addition to being used to initiate play by juveniles and adolescents, in the Taï South group (but not in the East group), the bending together of a few branches is used communicatively by adult males for sexual solicitation (e.g., Boesch, 2009, 2012b). Thus, the gesture in the South group is used in two contexts (each by one age-group) with two different meanings (invitation to play and sexual solicitation, see **Table 5**). It is unclear whether they evolved independently or whether one is based on the other. The latter would indicate a semantic shift as defined by Cissewski and Boesch (2016) and as described in the above section on semantic shifts, in this case limited to an age group. This would entail that the meaning of the gesture changed when the gesture was adopted by a different age group. This is visualized in **Figure 5**.

There are no systematic data on the use of MATING-NEST BUILDING. But as in the case of PLAY-NEST BUILDING, fieldworkers agree that the bent-over branches or saplings do not serve the original purpose of resting and that they exhibit communicative intent (e.g., Luncz, personal observation; Boesch, personal communication; Boesch, 2012a). MATING-NEST BUILDING is observed less frequently than PLAY-NEST BUILDING, because mating occurs less frequently than play and because MATING-NEST BUILDING (unlike PLAY-NEST BUILDING) is used only by male individuals.

MATING-NEST BUILDING consists of the quick bending together of a small number of branches or saplings, in close proximity to the female recipient and thus clearly audible and at least partly visible to her; it is done taking into account the potential mating partner's attentional state. The form of the resulting nests is usually simpler than that of day nests for resting, but they can get more elaborate if the recipient does not react quickly. The sender (the male) does not lie down on the branches after construction, he visually checks the attention of the recipient, and he exhibits response waiting. The sender receives a voluntary response, that is, the potential recipient (the female) is not pulled toward the sender. MATING-NEST BUILDING is thus mechanically ineffective toward the recipient. Due to reluctance of the potential mating partner, persistence or elaboration on the sender's side (e.g., by adding other context-specific signals) is often needed.

Thus, while PLAY-NEST BUILDING can often be observed as a "stand alone" gesture, MATING-NEST BUILDING typically occurs in connection with other context-specific signals like, for instance, the presenting of an erect penis. It also often becomes part of sequences of gestures with equivalent meaning (that is, sexual solicitation).

Video 3 shows a young chimpanzee male trying to convince a female (in the front with her back toward the camera) to mate with him. Due to her reluctance, the male uses a series of gestures, including MATING-NEST BUILDING at the beginning and, very rudimentary, in the middle of the sequence. The communicative intent of MATING-NEST BUILDING is nicely evident in the clip (e.g., monitoring the attention of the recipient, awaiting recipient's response).

Video 3: https://share.eva.mpg.de/index.php/s/AKH27jnrbLFy3Kp (Copyright: Liran Samuni, Taï Chimpanzee Project).

It is especially younger males who use this gesture (Luncz, personal observation), probably because more persuasion is necessary for an adult female to mate with them. The gesture here often seems to serve the purposes of persistence and elaboration, to persuade a female to accept the male's invitation. MATING-NEST BUILDING thus enlarges the number of gestures available for sexual solicitation and provides an additional means of persuading a female (especially an older female) to accept the male's invitation.[11]

In addition, MATING-NEST BUILDING is an inconspicuous means of signaling. In environments, where visibility is restricted and there is time pressure on signaling, it can be advantageous to have an inconspicuous signal that can attract the attention of a female situated within several meters, but not the attention of a dominant male further away. Moreover, the sender cannot be identified acoustically by distant group members. The lower-ranking males generally pay attention to not display the behavior in the vicinity of the alpha male so as to not be detected. Thus, by adopting the gesture, subordinate males may gain mating opportunities, and by reacting to it, females may gain choice of partners.

Crucially, as stated above, there is no logical connection between the building of proper nests and mating in the Taï South group. Real nests are not used for mating. Note further that the rudimentary construction that results from MATING-NEST BUILDING has no role in actual mating, because the attracted female approaches the sender and sexually presents outside the area of the construction.

---

[11] Note that also in human communication the use of additional signals and multimodality go on all the time. When agreeing, for instance, instead of simply saying "Yes." we often add other linguistic material, for instance "yes, that would be great." In addition, we may combine these words with affirmative gestures in different modalities like nodding (visual) or touching (tactile), or even stomping/clapping (auditory) to show extreme enthusiasm. In doing so intentionally or unintentionally, we provide the recipient(s) with information about our attitude and commitment.

Thus, the form of the gesture (bending a few saplings or branches together) and its meaning (sexual solicitation) would be linked arbitrarily (without a logical connection between the two) and the linkage would be conventionalized at the group-level. Note that MATING-NEST BUILDING (like PLAY-NEST BUILDING) has been observed for about two decades, showing its sustained use over generations and in this way providing evidence for acquisition *via* social learning. The reduced form of the resulting constructions further emphasizes the communicative intent of MATING-NEST BUILDING by making it distinguishable from the original underlying behavior (nest-building for resting). Given the points made above as well as the fact that MATING-NEST BUILDING can be used on a par in sequences with other established gestures (BRANCH SHAKING, PRESENTING PENIS), we propose that MATING-NEST BUILDING can be considered as a potential candidate for symbolic signal use for sexual solicitation in the Taï South group. However, more observational data are needed to establish whether MATING-NEST BUILDING is truly referential, that is, whether it is reliably understood by itself as a gesture for sexual solicitation. In contrast, PLAY-NEST BUILDING is reliably understood without further cues.

In summary, given that PLAY-NEST BUILDING and MATING-NEST BUILDING are group-specific gestures and thus cannot be found in other groups across the species, nor on the sub-species level, we would assume emergence and social transmission within an ontogenetic time frame, instead of innateness. The apparent arbitrary linkage between the gesture's form (bending together a small number of branches or saplings) and its meaning (play invitation/sexual solicitation) thus would be learned and in our opinion might constitute evidence for basic symbolic communication. However, systematic data collection is needed for the case to be conclusive.

Three Scenarios for the Emergence of the Ground-Nest Gesture in the Taï South Group

Based on the two processes of *emergence of semanticity* and *partial semantic shifts*, in this case study, we now apply these processes to the emergence of PLAY-NEST BUILDING and MATING-NEST BUILDING in the South group. We distinguish between two types of emergence: Independent emergence of PLAY-NEST BUILDING and MATING-NEST BUILDING in the two different age groups (scenario 1), and successive emergence, with MATING-NEST BUILDING (or PLAY-NEST BUILDING, respectively) being used first and then undergoing a partial semantic shift (scenarios 2–3) as suggested above and illustrated in **Figure 3**. An overview of the scenarios is provided in **Figure 5**.

CONCLUSION

Our aim is to contribute to our understanding of the developmental origins of great ape gestures and to inspire researchers studying wild primates (and other species) to systematically investigate group-specific gestures, and other learned communicative elements, against the background of potential symbolic signal use.

For identifying a basic form of symbolic signal use in great ape natural communication, we have provided a theoretical

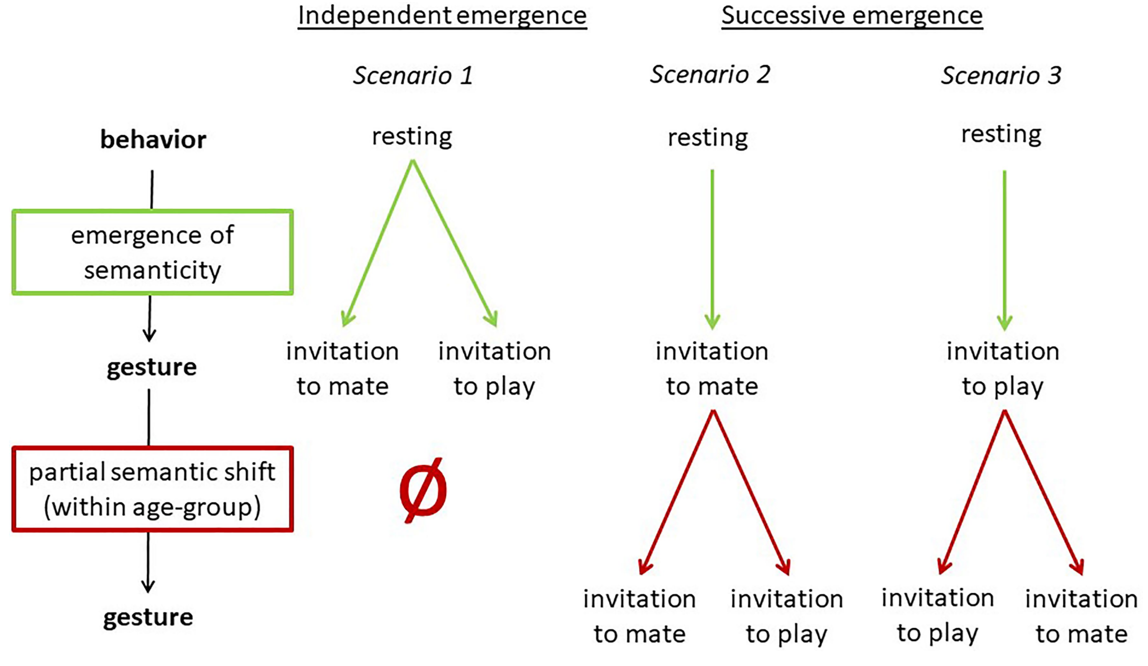


FIGURE 5 | Three scenarios for the emergence of PLAY-NEST BUILDING and MATING-NEST BUILDING in the Taï South group.

framework based on the key criteria of *arbitrariness* and *conventionalization*. The form-meaning linkage of a gesture thus would be *arbitrary* if there is no obvious logical or otherwise motivated connection between the form and the meaning. And the linkage would be *conventionalized* if the gesture is not innate but learned by the members of the respective groups.

For the rise of conventionalized arbitrary gestures, we have proposed two routes: *semantic shifts* (a change of meaning in an existing gesture) and the *emergence of semanticity* (the creation of new gestures on the basis of non-social behaviors). In both cases, the resulting gesture would exhibit an arbitrary linkage of form and meaning, because the form of the gesture was borrowed from a behavior outside the context in which the resulting gesture is used. This arbitrary linkage would be conventionalized at the group-level, resulting in a group-specific basic symbolic gesture.

Furthermore, we have suggested potential candidates for basic symbolism in chimpanzee natural gestural communication. These candidates seem to exhibit the key characteristics of symbolic signal use in our framework: an arbitrary and conventionalized form-meaning linkage. Compared with the symbolic capacities demonstrated by great apes in laboratory environments, these candidates suggest symbolic signal use in chimpanzee natural gestural communication to be rather limited both with regard to the number of possible candidates and with regard to the number of contexts. However, future systematic field research and analysis may reveal a richer picture both in number of examples and in their variety.

Our focus on the basic characteristics of symbolic signal use together with the suggestive data from the field shed new light on the existence, nature, and origin of chimpanzee symbolic gestural communication. By making the case for arbitrary and conventionalized signals to be accepted as a sufficient characteristic for the presence of basic symbolic signal use, we hope to widen the scientific perspective on symbolic communication across species boundaries and to contribute to a more complete assessment of the presence of symbolic gestures in our closest living relatives, the great apes.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JC wrote the first draft of this manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

Boesch, C. (1991). Symbolic communication in wild chimpanzees? *Hum. Evol.* 6, 81–89. doi: 10.1007/BF02435610

Boesch, C. (1995). Innovation in wild chimpanzees (*Pan troglodytes*). *Int. J. Primatol.* 16, 1–16. doi: 10.1007/BF02700150

Boesch, C. (2003). Is culture a golden barrier between human and chimpanzee? *Evol. Anthropol.* 12, 82–89. doi: 10.1002/evan.10106

Boesch, C. (2009). *The Real Chimpanzee: Sex Strategies in the Forest*. Cambridge: Cambridge University Press.

Boesch, C. (2012a). "From material to symbolic cultures: Culture in primates," in *The Oxford Handbook of Culture and Psychology*. ed. J. Valsiner (Oxford: Oxford University Press), 677–694.

Boesch, C. (2012b). *Wild Cultures: A Comparison Between Chimpanzee and Human Cultures*. Cambridge: Cambridge University Press.

Boesch, C., and Boesch-Achermann, H. (2000). *The Chimpanzees of the Taï Forest: Behavioural Ecology and Evolution*. Oxford: Oxford University Press.

Boesch, C., Crockford, C., Herbinger, I., Wittig, R. M., Möbius, Y., and Normand, E. (2008). Intergroup conflicts among chimpanzees in taï national park: lethal violence and the female perspective. *Am. J. Primatol.* 70, 519–532. doi: 10.1002/ajp.20524

Bonnie, K. E., and de Waal, F. B. M. (2006). Affiliation promotes the transmission of a social custom: handclasp grooming among captive chimpanzees. *Primates* 47, 27–34. doi: 10.1007/s10329-005-0141-0

Byrne, R. W., Cartmill, E. A., Genty, E., Graham, K. E., Hobaiter, C., and Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Anim. Cogn.* 20, 755–769. doi: 10.1007/s10071-017-1096-4

Cartmill, E. A., and Hobaiter, C. (2019). Gesturing towards the future: cognition, big data, and the future of comparative gesture research. *Anim. Cogn.* 22, 597–604. doi: 10.1007/s10071-019-01278-x

Chandler, D. (2017). *Semiotics: The Basics. 3rd Edn*. London: Routledge.

Cheney, D. L., and Seyfarth, R. M. (2007). *Baboon Metaphysics*. Chicago: The University of Chicago Press.

Cissewski, J., and Boesch, C. (2016). Communication without language: how great apes may cover crucial advantages of language without creating a system of symbolic communication. *Gesture* 15, 224–249. doi: 10.1075/gest.15.2.04cis

Crockford, C., Herbinger, I., Vigilant, L., and Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* 110, 221–243. doi: 10.1111/j.1439-0310.2004.00968.x

Crystal, D. (2019). *The Cambridge Encyclopedia of the English Language. 3rd Edn*. Cambridge: Cambridge University Press.

Darwin, C. (1872). *The Expression of Emotion in Man and Animals*. London: John Murray.

Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. London: Penguin.

Deacon, T. W. (2012). "Beyond the symbolic species," in *The Symbolic Species Evolved*. eds. T. Schilhab, F. Stjernfelt and T. Deacon (Heidelberg London, New York: Springer), 9–38.

de Saussure, F. (1916). *Cours de linguistique générale*. eds. C. Bally, A. Séchehaye and A. Riedlinger (Paris: Payot).

Fruth, B., and Hohmann, G. (1996). "Nest building behavior in the great apes: the great leap forward?" in *Great Ape Societies*. eds. W. C. McGrew, L. F. Marchant and T. Nishida (Cambridge: Cambridge University Press), 225–240.

Genty, E., Breuer, T., Hobaiter, C., and Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim. Cogn.* 12, 527–546. doi: 10.1007/s10071-009-0213-4

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: The Belknap Press of Harvard University Press.

Graham, K. E., Furuichi, T., and Byrne, R. W. (2017). The gestural repertoire of the wild bonobo (*Pan paniscus*): a mutually understood communication system. *Anim. Cogn.* 20, 171–177. doi: 10.1007/s10071-016-1035-9

Greenfield, P. M., and Savage-Rumbaugh, E. S. (1990). "Grammatical combination in *Pan paniscus*: processes of learning and invention in the evolution and development of language," in *Language' and Intelligence in Monkeys and Apes: Comparative Developmental Perspectives*. eds. S. T. Parker and K. R. Gibson (New York: Cambridge University Press), 540–578.

Halina, M., Rossano, F., and Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Anim. Cogn.* 16, 653–666. doi: 10.1007/s10071-013-0601-7

Herman, I. M. (1987). "Receptive competencies of language-trained animals," in *Advances in the Study of Behavior*. *Vol. 17*. eds. J. S. Rosenblatt, C. Beer, M.-C. Busnel and P. J. B. Slater (San Diego, CA: Academic Press), 1–60.

Hicks, T. C. (2010). A Chimpanzee Mega-Culture? Exploring behavioral continuity in Pan troglodytes schweinfurthii across northern DR Congo. Doctoral dissertation. University of Amsterdam.

Hobaiter, C., and Byrne, R. W. (2010). Able-bodied wild chimpanzees imitate a motor procedure used by a disabled individual to overcome handicap. *PLoS One* 5:e11959. doi: 10.1371/journal.pone.0011959

Hobaiter, C., and Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Anim. Cogn.* 14, 745–767. doi: 10.1007/s10071-011-0409-2

Hobaiter, C., and Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Curr. Biol.* 24, 1596–1600. doi: 10.1016/j.cub.2014.05.066

Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 88–96. doi: 10.1038/scientificamerican0960-88

Hoppitt, W. J. E., and Laland, K. N. (2008). Social processes influencing learning in animals: a review of the evidence. *Adv. Study Behav.* 38, 105–165. doi: 10.1016/s0065-3454(08)00003-x

Hoppitt, W. J. E., and Laland, K. N. (2013). *Social Learning: An Introduction to Mechanisms, Methods and Models*. Princeton: Princeton University Press.

Kalan, A. K., and Boesch, C. (2018). Re-emergence of the leaf clip gesture during an alpha takeover affects variation in male chimpanzee loud calls. *PeerJ* 6:e5079. doi: 10.7717/peerj.5079

Kaminski, J., Call, J., and Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science* 304, 1682–1683. doi: 10.1126/science.1097859

Krebs, J. R., and Dawkins, R. (1984). "Animal signals: mind-reading and manipulation," in *Behavioral Ecology: An Evolutionary Approach*. *2nd Edn.* eds. J. R. Krebs and N. B. Davies (Oxford: Blackwell Scientific Publications), 380–402.

Liebal, K., and Call, J. (2012). The origins of non-human primates' manual gestures. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 118–128. doi: 10.1098/rstb.2011.0044

Liebal, K., Pika, S., Call, J., and Tomasello, M. (2004). To move or not to move: how great apes adjust to the attentional state of others. *Interact. Stud.* 5, 199–219. doi: 10.1075/is.5.2.03lie

Liebal, K., Pika, S., and Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6, 1–38. doi: 10.1075/gest.6.1.02lie

Liebal, K., Schneider, C., and Errson-Lembeck, M. (2019). How primates acquire their gestures: evaluating current theories and evidence. *Anim. Cogn.* 22, 473–486. doi: 10.1007/s10071-018-1187-x

Luncz, L. V., and Boesch, C. (2014). Tradition over trend: neighboring chimpanzee communities maintain differences in cultural behavior despite frequent immigration of adult females. *Am. J. Primatol.* 76, 649–657. doi: 10.1002/ajp.22259

Luncz, L. V., and Boesch, C. (2015). The extent of cultural variation between adjacent chimpanzee (*Pan troglodytes* verus) communities: a microecological approach. *Am. J. Phys. Anthropol.* 156, 67–75. doi: 10.1002/ajpa.22628

Luncz, L. V., Mundry, R., and Boesch, C. (2012). Evidence for cultural differences between neighboring chimpanzee communities. *Curr. Biol.* 22, 922–926. doi: 10.1016/j.cub.2012.03.031

MacKinnon, J. R. (1974). The behaviour and ecology of wild orang-utans (*Pongo pygmaeus*). *Anim. Behav.* 22, 3–74. doi: 10.1016/S0003-3472(74)80054-0

Miles, H. L. (1990). "The cognitive foundations for reference in a signing orangutan," in *Language and Intelligence in Monkeys and Apes: Comparative Developmental Perspectives*. eds. S. T. Parker and K. R. Gibson (Cambridge: Cambridge University Press), 511–539.

Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate mental development. *Am. Anthropol.* 83, 534–548. doi: 10.1525/aa.1981.83.3.02a00020

Nishida, T. (1980). The leaf-clipping display: a newly-discovered expressive gesture in wild chimpanzees. *J. Hum. Evol.* 9, 117–128. doi: 10.1016/0047-2484(80)90068-8

Nishida, T. (1987). "Local traditions and cultural transmission," in *Primate Societies*. eds. B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham and T. T. Struhsaker (Chicago: Chicago University Press), 462–474.

Patterson, F. G. (1978). The gestures of a gorilla: language acquisition in another Pongid. *Brain Lang.* 5, 72–97. doi: 10.1016/0093-934X(78)90008-1

Pepperberg, I. M. (2017). "Symbolic communication in nonhuman animals," in *APA Handbook of Comparative Psychology: Basic Concepts, Methods, Neural Substrate, and Behaviour*. eds. J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon and T. Zentall (Washington, DC: American Psychological Association), 663–679.

Pepperberg, I. M., and Nakayama, K. (2016). Robust representation of shape by a grey parrot (*Psittacus erithacus*). *Cognition* 153, 146–160. doi: 10.1016/j.cognition.2016.04.014

Pika, S., Liebal, K., and Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): gestural repertoire, learning, and use. *Am. J. Primatol.* 60, 95–111. doi: 10.1002/ajp.10097

Pika, S., Liebal, K., and Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): repertoire and use. *Am. J. Primatol.* 65, 39–61. doi: 10.1002/ajp.20096

Russon, A. E., and Begun, D. R. (2004). "Evolutionary origins of great ape intelligence: an integrated view," in *The Evolution of Thought*. eds. A. E. Russon and D. R. Begun (Cambridge: Cambridge University Press), 353–368.

Samuni, L., Preis, A., Mundry, R., Deschner, T., Crockford, C., and Wittig, R. M. (2017). Oxytocin reactivity during intergroup conflict in wild chimpanzees. *Proc. Natl. Acad. Sci. U. S. A.* 114, 268–273. doi: 10.1073/pnas.1616812114

Savage-Rumbaugh, E. S., McDonald, K., Sevcik, R. A., Hopkins, W. D., and Rubert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *J. Exp. Psychol. Gen.* 115, 211–235. doi: 10.1037/0096-3445.115.3.211

Schubert, G., Stoneking, C. J., Arandjelovic, M., Boesch, C., Eckhardt, N., Hohmann, G., et al. (2011). Male-mediated gene flow in patrilocal primates. *PLoS One* 6:e21514. doi: 10.1371/journal.pone.0021514

Schusterman, R., and Krieger, K. (1984). California sea lions are capable of semantic comprehension. *Psychol. Rec.* 34, 3–23. doi: 10.1007/BF03394849

Sugiyama, Y. (1981). Observations on the population dynamics and behavior of wild chimpanzees at Bossou, Guinea, in 1979–1980. *Primates* 22, 435–444. doi: 10.1007/BF02381236

Tinbergen, N. (1952). "Derived" activities; their causation, biological significance, origin, and emancipation during evolution. *Q. Rev. Biol.* 27, 1–32. doi: 10.1086/398642

Tomasello, M. (1996). "Do apes ape?" in *Social Learning in Animals: The Roots of Culture*. eds. J. Galef and C. Heyes (Academic Press), 319–346.

Tomasello, M., and Call, J. (2007). "Intentional communication in nonhuman primates," in *The Gestural Communication of Apes and Monkeys*. eds. J. Call and M. Tomasello (Mawah: Lawrence Erlbaum Associates), 1–15.

Tomasello, M., and Call, J. (2019). Thirty years of great ape gestures. *Anim. Cogn.* 22, 461–469. doi: 10.1007/s10071-018-1167-1

Tomasello, M., Gust, D., and Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates* 30, 35–50. doi: 10.1007/BF02381209

van Hooff, J. A. (1972). "A comparative approach to the phylogeny of laughter and smiling," in *Non-verbal Communication*. ed. R. Hinder (Cambridge: Cambridge University Press), 209–237.

van Hooff, J. A. (2012). "Facial-vocal displays, gestures and language," in *Developments in Primate Gesture Research*. eds. S. Pika and K. Liebal (Amsterdam/Philadelphia: John Benjamins Publishing Company), 13–31.

van Leeuwen, E. J. C., Cronin, K. A., and Haun, D. B. M. (2014). A group-specific arbitrary tradition in chimpanzees (*Pan troglodytes*). *Anim. Cogn.* 17, 1421–1425. doi: 10.1007/s10071-014-0766-8

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., et al. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Curr. Biol.* 25, 495–499. doi: 10.1016/j.cub.2014.12.032

Watts, D. P. (2008). Tool use by chimpanzees at Ngogo, Kibale National Park, Uganda. *Int. J. Primatol.* 29, 83–94. doi: 10.1007/s10764-007-9227-4

Webster, A. (2017). "Reading, writing and thinking," in *Deafness, Development and Literacy*. ed. A. Webster (London: Routledge), 148–206.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature* 399, 682–685. doi: 10.1038/21415

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (2001). Charting cultural variation in chimpanzees. *Behaviour* 138, 1481–1516. doi: 10.1163/156853901317367717

Whiten, A., Horner, V., Litchfield, C. A., and Marshall-Pescini, S. (2004). How do apes ape? *Anim. Learn. Behav.* 32, 36–52. doi: 10.3758/BF03196005

Wich, S. A., Krützen, M., Lameira, A. R., Nater, A., Arora, N., Bastian, M. L., et al. (2012). Call cultures in orang-utans? *PLoS One* 7:e36180. doi: 10.1371/journal.pone.0036180

Wilson, E. O. (1975). *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard University Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

HYPOTHESIS AND THEORY
published: 14 January 2022
doi: 10.3389/fpsyg.2021.718251

Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates

Elizabeth Warren* and Josep Call

School of Psychology and Neuroscience, University of St Andrews, St. Andrews, United Kingdom

OPEN ACCESS

Edited by:
Ulrike Griebel,
University of Memphis, United States

Reviewed by:
Brandon Wheeler,
University of Kent, United Kingdom
Nathalie Gontier,
University of Lisbon, Portugal

****Correspondence:***
Elizabeth Warren
ew84@st-andrews.ac.uk

Specialty section:
This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: *31 May 2021*
Accepted: *13 December 2021*
Published: *14 January 2022*

Citation:
Warren E and Call J (2022)
Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates.
Front. Psychol. 12:718251.
doi: 10.3389/fpsyg.2021.718251

Communication, when defined as an act intended to affect the psychological state of another individual, demands the use of inference. Either the signaler, the recipient, or both must make leaps of understanding which surpass the semantic information available and draw from pragmatic clues to fully imbue and interpret meaning. While research into human communication and the evolution of language has long been comfortable with mentalistic interpretations of communicative exchanges, including rich attributions of mental state, research into animal communication has balked at theoretical models which describe mentalized cognitive mechanisms. We submit a new theoretical perspective on animal communication: the model of inferential communication. For use when existing proximate models of animal communication are not sufficient to fully explain the complex, flexible, and intentional communication documented in certain species, specifically non-human primates, we present our model as a bridge between shallower, less cognitive descriptions of communicative behavior and the perhaps otherwise inaccessible mentalistic interpretations of communication found in theoretical considerations of human language. Inferential communication is a framework that builds on existing evidence of referentiality, intentionality, and social inference in primates. It allows that they might be capable of applying social inferences to a communicative setting, which could explain some of the cognitive processes that enable the complexity and flexibility of primate communication systems. While historical models of animal communication focus on the means-ends process of behavior and apparent cognitive outcomes, inferential communication invites consideration of the mentalistic processes that must underlie those outcomes. We propose a mentalized approach to questions, investigations, and interpretations of non-human primate communication. We include an overview of both ultimate and proximate models of animal communication, which contextualize the role and utility of our inferential communication model, and provide a detailed breakdown of the possible levels of cognitive complexity which could be investigated using this framework. Finally, we present some possible applications of inferential communication in the field of non-human primate communication and highlight the role it could play in advancing progress toward an increasingly precise understanding of the cognitive capabilities of our closest living relatives.

Keywords: animal communication, primates, social inference, communication cognition, intentionality, inferential communication, cognitive flexibility, social cognition

INTRODUCTION

Communication modifies the behavior of others by altering the psychological state of the recipient. Unlike instrumental actions, which bypass the recipient's psychological states and act directly on their behavior, communicative acts affect the perception, attention and/or cognition of recipients, and, if successful, subsequently provoke the desired behavior. Consider, for example, an infant chimpanzee who, while clinging to their mother, begins to nurse. The infant is engaged in an instrumental action with a direct effect on the mother's body, without engagement with the mother's psychological state. Although the mother could choose to disrupt the infant's feeding behavior if she did not desire that interaction, the feeding interaction itself is instrumental, not communicative. Conversely, an infant chimpanzee who reaches their hand toward their mother's back, a ritualized gesture which requests carrying (Hobaiter and Byrne, 2014), is altering the mental state of the mother, who may react to her perception and cognitive processing of this event by lifting the infant onto her back and performing the desired carrying behavior. Although the ultimate outcomes of the two interactions are similar – the infant's physical needs are met – the proximate mechanisms that permitted these outcomes are fundamentally different. The proximate mechanisms of communication, the alteration of psychological states to influence behavior, are an exceptional lens through which we can probe the levels of cognitive engagement involved in different communication systems.

Psychological states play a central role in all forms of communication, from the wing spots of a butterfly to the courtship display of a gull to linguistic exchanges between humans. These systems of communication differ, however, in their origins and, more importantly for our purposes here, in fixedness of the signals and in how likely they trigger certain responses in the audience that receive them. In cases where invariable signals precede invariable responses, there is little room for cognition. Therefore, dispensing of the cognitive 'waystation' in such cases does not represent a substantial loss, and communication can be viewed as signals or actions used to alter behavior. The breadth of communicative behavior, however, cannot be fully encompassed by fixed signals with involuntary responses.

Bypassing cognition becomes more difficult when the signals and responses are not fixed, but rather show some degree of variability. Flexibility in communication was first recognized by zoosemioticians studying the meaning of animal signals (e.g., Marler, 1961; Plooij and Lock, 1978), and later by researchers interested in intentional and goal-directed communication (e.g., Tomasello et al., 1985; Byrne et al., 2017). Both the early "signal meaning" approaches and the later intentional/goal-directed approaches to communication address cognitive aspects, but we will argue that neither of them are sufficient to fully explore how animals might use psychological states, and particularly some forms of inference about mental states, to communicate. In fact, some recent contributions that have embraced cognitive models of communication (e.g., Townsend et al., 2017) have flatly rejected mentalizing at any level and instead focus on superficial features of communication that denote flexible cognition. We think that this is a regressive mistake. The Gricean approach (Grice, 1957, 1969), which theorizes a high level of cognitive complexity, including pragmatic meaning, in communicative exchanges, is difficult to implement in investigations and interpretations of animal communication. The central idea that mentalizing plays a role in animal communication, however, deserves careful consideration.

One problem with completely rejecting mentalizing in animal communication, particularly if one is interested in the flexibility of a communicative system, is that mentalizing unlocks an unprecedented level of flexibility in human communication. Since many cognitive approaches to animal communication have used human communication as a point of comparison, particularly in considerations of the evolutionary origins of human language (e.g., Hewes et al., 1973; Zuberbühler, 2005; Scott-Phillips, 2015), it is at the very least questionable to *a priori* discard mentalizing. Although documenting flexibility in animal communication by means of behavioral indicators such as means-ends dissociations, contextual variation in signal use, and audience effects is a necessary first step (e.g., Tomasello and Zuberbühler, 2002; Tomasello, 2009 for review), such indicators explain neither the origin nor the psychological underpinnings of flexible responses. Producing a descriptive list of behavioral indicators of flexibility (goal-directedness) without digging deeper into the psychological process that give rise to those responses seems a missed opportunity. The problem is further compounded by the fact that referential and intentional communication are often used to explain language evolution (e.g., Arbib et al., 2008), but language is a system with mentalizing at its core (Grice, 1957, 1969; Wilson and Sperber, 2002). Without postulating some ability to make inferences about mental state to some forms of animal communication, the leap from animal to human communication, and language in particular, might be too great to be realistic. If the cognitive complexity of human communication is the measuring stick against which animal systems of communication are compared, at least in investigations of the evolutionary origins of language, then there exists a gap between the complexity of communicative behavior explained by the intentional model of communication and the ostensive-inferential models of human communication whose potential application has been discussed in certain animals, such as non-human primates (hereafter, "primates").

In this article we propose a solution to this gap in current models' explanatory power, for use in situations where the communicative behavior of a species or taxa involves an apparent level of flexibility and pragmatism not fully explained by existing models. We would like to introduce a model of communication – "inferential communication" – which we will distinguish from the model of intentional communication (e.g., Woodruff and Premack, 1979) and differentiate from other descriptions of inferential communication discussed by Fischer (2013); Fitch (2015) as well as those posited in developmental literature and studies of linguistics (e.g., Sperber and Wilson, 1986; Scott-Phillips, 2015; Moore, 2016). Our model is not intended encompass the same scope as global models of animal communication with ultimate explanations for communicative behavior. We submit inferential communication as a proximate model of communication which elaborates on ultimate explanations of communicative behavior by outlining some of the cognitive mechanisms that may operate within these ultimate models.

As we hope will become apparent, our proposal differs from cognitive models of animal communication that incorporate inferential processes on the one hand, and human ostensive communication on the other, along three main dimensions: the nature of the inference, the type of pragmatics involved and the role of informative intentions. We will combine the comparative research that has been accumulated in the last three decades on referential and intentional communication with data on social cognition and inferential reasoning to establish the theoretical foundations for our perspective on inferential communication. Thus, one of our key proposals is that mental state attribution, rather than being a problem, it is part of the solution. Together with inferential reasoning, it constitutes the cognitive substrate of flexible communication.

Our paper is organized as follows. First, we will summarize the traditional approaches to animal communication, in order of increasing engagement with cognition, and provide the theoretical background to contextualize the model we now propose. Second, we will outline the model of inferential communication, specifically with respect to primates, distinguishing our proposal from previous characterizations of inference in communication. Third, we will delineate the cognitive skills and mechanisms required for each increasingly mentalized level of complexity within our model. Fourth, we will shed light on the applications of inferential communication, from both a theoretical and experimental perspective, and explain the breadth of taxa to which it could potentially be applied. Finally, we will place inferential communication into the broader field of theoretical approaches to primate communication.

FROM SIGNALS TO INTENTIONAL GESTURES

Manipulation Model

To appreciate the theoretical justification for inferential communication, it is critical to review both the tenets of ultimate approaches to animal communication and the questions they leave unanswered. The earliest ethological models of animal communication, including non-human primate communication, were founded in behavior, not cognition (see **Table 1**). Building on the work of Tinbergen (1952), Lorenz (1966), who created the foundation for phylogenetic preservation of evolutionarily successful behaviors, Dawkins and Krebs (1978), Krebs and Dawkins (1984) asserted that animal systems of communication are the result of repeated, non-communicative instrumental actions that become phylogenetically ritualized to prompt certain behavioral responses in others. Just as instrumental actions affect the environment to produce a certain result, communicative signals act on others to induce certain behaviors. If successful, the signaler will have incurred benefit as a result of the exchange, and thus the signal persists as a function of evolutionary fitness. This non-mentalized, behavior-centric approach is upheld in some modern work (e.g., Owren et al., 2010), where animal communication is described as an effort to influence the behavior of another and is placed in the shared evolutionary timeline of living primate species, including humans, as a necessary but distant step in the evolution of human language.

If we apply this model to an example of a communicative interaction between two primates, the ritualized format of the exchange becomes clearer. In this example, one primate, Cindy wishes to be groomed by another primate, Louis. Accordingly, Cindy moves toward Louis and presents her shoulder, a behavioral pattern known to culminate in the receipt of grooming (Hobaiter and Byrne, 2014). Louis grooms Cindy's shoulder, and Cindy therefore receives fitness benefits associated with grooming. Viewing this exchange through the lens of communication as manipulation, Cindy has engaged in a ritualized action which likely developed from the necessary instrumental actions associated with grooming, i.e., moving the body part close enough to allow grooming to occur. This action manipulated a response from Louis, the outcome of which benefited Cindy, who is therefore likely to repeat the gesture in the future, and the gesture is maintained, over evolutionary time, in this primate gestural repertoire. Notably, the ritualization of gestures here is from a phylogenetical perspective, not an individual one, and thus does not ascribe an individual representation or any cognitive process underlying the behavior to either party.

This model of communication offers an ultimate explanation of communication with broad taxonomic applicability; the same principles of manipulation and evolutionary fitness that explain the phylogenetic preservation of primate gestures explain the mating display of a bower bird or the aposematism of a toxic insect. This model does not, however, offer proximate explanations for the behavioral patterns of communication; it allows for situations where the induced response of the recipient is the result of *understanding* the manipulation and situations where the induced response is merely *a reaction to* the manipulation, the latter of which requires no cognitive engagement with, or even awareness of, the signaler's desired outcome. There is an opportunity, therefore, for proximate models of communication to elaborate on the means-ends process of communication-as-manipulation by positing the mechanisms that might underlie the communicative behaviors.

Information Model

Following Shannon and Weaver (1949), Marler (1961) proposed the theory of animal communication as information. This model characterizes information as the reduction of uncertainty on the part of the recipient, where the signaler encodes signals with informational meaning, and the recipient can decode these signals to access information. Although the informative signals are not necessarily under the intentional control of the signaler, they are still adaptive, just as in the manipulation model in the sense that they facilitate the desired outcome from the recipient. As a complement to the manipulation model, which more readily explains the fitness benefit of the signaler, the information model explains the adaptive benefit to the recipient more clearly – the recipient can achieve greater fitness by properly decoding the signal, gaining easier access to cooperative, affiliative exchanges, as well as easier interpretation of fearful, aggressive, or competitive displays.

Although the information model, which predates the manipulation model (see Dawkins and Krebs, 1978; Krebs and Dawkins, 1984), is still mainly centered on an ultimate perspective on communication, its principles eventually facilitated research on the cognitive mechanisms underlying communication. The process of giving and receiving informative signals can involve cognitive skills, including semantic encoding/decoding and functional reference (see **Table 1**). Furthermore, within this informational model, signals cannot necessarily be mapped 1-1 onto meanings, but may demand the use of contextual cues for accurate decoding (Smith, 1977).

Following our earlier example of an exchange between primates, the informational model of communication would interpret the actions *via* the route of informational transmission. Cindy wants to be groomed, and she encodes this information in a signal – a big, loud scratch across her own chest (Hobaiter and Byrne, 2014). As the scratching behavior is a non-instrumental signal, meaning that it does not act directly on the body of the recipient, Louis must decode this signal based on contextual cues and existing knowledge of the signal, and in doing so, receives the information that Cindy wants to be groomed. Louis may produce the desired behavior, or not, depending on the context and the fitness benefit to himself. Not only does this informational perspective address the success of the exchange from the perspective of both the signaler and the recipient, it also opens the door for an element of cognition: encoding and decoding of non-instrumental signals. Although not all informative signals require encoding and decoding – it is equally possible to inadvertently signal information and induce an innate reaction to that information – encoding and decoding become possible under this model of communication, which permits questions relating to cognitive engagement with the act of communication.

TABLE 1 | Ethological models of communication including the origin and signal-referent relation as well as their key cognitive concepts.

Model	Sub-discipline	Signal origin	Signal-referent relation	Key cognitive concepts
Manipulation	Behavioral ecology	Innate	Fixed	n/a
Information	Zoosemiotics	Innate/learned	Flexible	Semantic signal encoding and decoding functional reference

The informational perspective, though more robust in its mechanistic considerations, is more a behavioral model of communication than a cognitive one, and thus has theoretical limitations in its ability to fully characterize the cognitive abilities of certain species within communication. It describes cognitive engagement on the level of signal decoding and introduces the concept of flexible interpretation (i.e., varied interpretation of the same signal based on context). It does not, however, address the question of referentiality, at least, not in its earlier iterations (Seyfarth et al., 1980). Vocalizations or gestures encoded with information could be produced voluntarily or involuntarily, while still consisting of a non-instrumental signal encoded with valuable information for the recipient. Modern work within this paradigm (e.g., Tomasello and Zuberbühler, 2002; Leavens et al., 2004), asks this question of intentionality and referentiality, but does not conclusively conclude that the signaler or the recipient have an internal representation of the information, and rather, could be exhibiting "functional referentiality," characterized by signals provoked directly by the external stimuli about which they contain information (Slocombe and Zuberbühler, 2005). Without an ability to account for internal representation of intention and meaning, the informational model of communication is inherently limited to basic, practicable cognitive mechanisms – encoding and decoding – which do not encompass the rich breadth of possible mentalizing in primate communication.

INTENTIONAL COMMUNICATION

Intentional communication, also known as goal-directed communication, the third and final historical model of communication, can be considered the first of three fully cognitive models (see **Table 2**). It introduced two critical cognitive skills – intentionality and goal-directed signals. Plooij and Lock (1978), Woodruff and Premack (1979) were among the first to thoroughly address the question of intentionality in animal communication, specifically in the communication system of primates. They characterized intentional communication as transmission of information between a signaler and a recipient adhering to three main criteria: first, the signaler must be aware the transmission of information will result from the signal; second, the signaler expects that the recipient will similarly be aware of the

TABLE 2 | Psychological models of animal communication including the signal origins, the signaler's intention, the recipient's decoding, and the cumulative requisite cognitive skills (later models include those of previous ones).

Model	Signal origin	Signaler's intention	Recipient's inference	Cognitive skills
Intentional Communication	Phylogenetic ritualization Ontogenetic ritualization	I want her to do X for me	n/a (I will do X to her)	Goal-directed signals Intentionality Referentiality Awareness of informational transmission
Inferential Communication	Inference	I want her to do X(= $\mathbf{x_1} + \mathbf{x_2} + \mathbf{x_3}$) for me	**What does** she want me to do to her**?**	Prosociality Informative intention
Ostensive Communication	Conventionalization Imitative learning	I want **to tell** her to do X for me	What does she want **to tell** me to do to her?	Communicative intention Recursive mental states/3rd- and 4th-order theory of mind

A key aspect of inferential communication is that the signaler creates a new signal (or modifies an existing one) to instruct the recipient what to do. X(= $\mathbf{x_1} + \mathbf{x_2} + \mathbf{x_3}$*) is meant to indicate that the signaler provides not just information about their goal, but also instruction about how to do a particular action. Bold lettering in the signaler and recipient column indicates the new component in each model compared to the previous one.*

transmission of information; and finally, the signaler must be able to selectively control their own signals in order to transmit the desired information. Later work (e.g., Tomasello et al., 1985, 1989; Hopkins et al., 2007; Byrne et al., 2017) on intentional communication follows several core criteria for intentionality, first defined by Bretherton and Bates (1979) for use in developmental psychology. These core hallmarks of intentionality include attentional monitoring, gaze-alternation, persistence, and elaboration.

While at least a subset of these criteria are necessary to indicate intentionality, they alone are not sufficient to conclusively demonstrate it. Townsend et al. (2017) note that, although there is no specific combination of criteria that would absolutely indicate intentionality, more indicators for any particular species or experiment serve as stronger evidence that the intentionality is genuine. Furthermore, we argue that intentionality is most likely to be at work when it is robust in the face of experimental perturbation. If flexible, apparently intentional communication cannot be transferred to a new situation where the old conditions of the successful communicative exchanges do not apply, and exchanges are unsuccessful in this new setting, then the communicative system may be more rigid than initially indicated by successful demonstration of the above criteria. Vail et al. (2013) demonstrated several attributes of intentional communication in coral reef fish (*Plectropomus pessuliferus marisrubri*), theoretically suggesting that intentionality may be more widespread than the complexity of the behavior might suggest. It is unknown, however, whether the apparently referential signals in fish would stand up under multiple, varied circumstances, which would be stronger evidence of flexible, goal-directed, intentional communication. If it was indeed the case that coral reef fish could successfully transfer this behavior to a new situation, then there would be no reason to deny the potential for intentional communication in their species. Each of the criteria for intentionality, including flexible transference of the intentionality to new circumstances, has been demonstrated, experimentally or observationally, in primates, particularly great apes (Leavens et al., 2005 for review; Graham et al., 2020).

Carrying our primate grooming example forward, we now apply the intentional model of communication to these actions. Cindy, the signaler, must first open an attentional channel with Louis, the recipient, ensuring that she has his attention either through the use of an auditory or tactile "attention-getter" signal (Leavens et al., 2005), or by checking for existing visual contact. Cindy must have an internal representation of what she wants – grooming – and an awareness that she needs to transmit information about her goal – her desire for grooming – to Louis. She produces the signal, the big loud scratch from earlier, intentionally, and monitors Louis' response, to determine whether the communication was sufficient to meet her internally represented goal. Louis, the recipient, must attend to Cindy, and must be aware that information is encoded in the signal, thus prompting him to decode it. As before, Louis can provide the desired grooming

FIGURE 1 | Illustration of two primates engaged in a communicative exchange depicting the recipient's (lack of) inference under the intentional communication model. Illustration by Sadie Tenpas.

behavior, or not, at which point Cindy may persist and produce the gesture again, or she may elaborate, by producing a different signal which can also be decoded to request grooming (see **Figure 1**).

Cindy, in this example, is displaying new cognitive skills compared to those demanded by the previous models of communication. She is engaging in goal-directed communication, where she is internally motivated by her own goal and is using communication as a means of achieving it. She is displaying intentionality, wherein her actions are under her voluntary control, and, at this stage, she is communicating referentially, in that she is making direct, intentional reference to what she wants. Notably, the same cognitive mechanisms are not necessarily required of Louis, other than conscious awareness of the transmission of information. Although, according to this model of communication, he must be aware that there is information to decode, his response to that information does not necessarily need to be voluntary, for the communication to be successful. As in above examples, his response to the information he has decoded could be innate, or externally motivated by the stimulus of the information, rather than motivated by his own internal representation of Cindy's goal.

Research using the framework of intentional communication has amassed a substantial body of evidence to support flexibility in primate communication (Liebal et al., 2014, for review). Regarding referentiality in primates, several studies have found evidence to support functional referentiality in the vocal domain (e.g., Slocombe and Zuberbühler, 2005, but see Fischer and Price, 2017 for an opposing view), and in the gestural domain (Call and Tomasello, 2007 for review). These are crucial findings for intentional models of communication, and they provide a framework within which to describe some of the flexible and behaviorally complex communication observed in primates from a cognitive standpoint. In our view, however, they still fall short of fully explaining the mechanisms at play in production and interpretation of communicative behavior in species with complex cognitive engagement during communicative acts. Intentional communication, as a model, invokes a means-ends dissociation, in that it describes observed behaviors in the context of their relevant psychological effects, but does not delve into the actual cognitive processes that permit these cognitive outcomes. It is clear that intentionality and flexibility place cognitive demands on both the signaler and the recipient, but the exact psychological processes are not illuminated. In fact, when we thought that the field was ripe to explore the psychological processes in greater detail, researchers have hesitated to take what we see as the next necessary step in unraveling the complexity of primate communication. In order to further advance our understanding of the cognitive mechanisms underpinning communication, we need to look beyond intentionality and toward psychological states. We propose the model of inferential communication as a means of explaining and investigating the cognitive, mentalistic aspects of communication, and to form a bridge between existing models of primate communication and the ostensive, language-oriented models found in the human developmental literature.

THE MODEL OF INFERENTIAL COMMUNICATION

As a theoretical model, inferential communication can be viewed as a system of conveying messages which operates outside the confines of codified, semantic gestures or vocalizations (Wilson, 1998), and which requires the integration of known information and context to interpret informational meaning. While we do not assert that inferential communication is engaged during all communicative interactions in any species, including humans, we submit this model as an explanatory and heuristic tool to investigate communicative behavior where inferential leaps of understanding, for both signalers and recipients, are required for successful transmission of information. When alternative explanations of apparently successful communicative behavior are ruled out, it allows for the investigation of higher-order cognitive mechanisms, such as mental state interpretation, prosociality, and, most crucially, *rational inference*. Crucially, in our model, inferential thinking is required of both the signaler, who must account for the leaps of understanding the recipient may make when deciding on the level of ambiguity in the signal, and the recipient, who must infer the meaning of the information being conveyed.

Many authors have noted there is ample evidence that recipients infer meaning from signals (Fischer, 2013; Fitch, 2015; Fischer and Price, 2017; Seyfarth and Cheney, 2017). However, the kind of inferred meaning that we endeavor to investigate differs from other proposals in terms of the type of inference that supports the communication and the type of pragmatics involved. First, we use inference more narrowly than other authors, to distinguish it from other processes. In a broad sense, when a baboon hears the call of his consort behind some bushes, he may infer that she is located behind those bushes (Fischer, 2013). But it is also possible that the individual has learned over time that when that call is produced, a particular female will appear behind those bushes – so an association rather than an inference might be doing the work of deciphering the signal. Another interpretation of "inference" refers the integration of information from multiple sources to make a decision (Fitch, 2015; Fischer and Price, 2017). There is no doubt that integration – putting together disparate pieces of information - is a fundamental aspect of inference (Tolman, 1932; Premack and Premack, 1994). But integration can also be achieved by processes such as conditional discrimination. When a baboon hears a specific female's call, but he also sees that her juvenile offspring are nearby, he may respond differently to her call than if they were absent, not because he has inferred different meaning from her signal, but because he has learned over repeated exposure to similar situations that the appropriate response differs from a situation in which he is alone with the female. In this case, he is not exhibiting inference, but merely learned different responses to different contexts. We agree that inference requires the integration and assimilation of multiple pieces of information to guess at outcomes (i.e., "contextual pragmatics" in Fischer and Price, 2017). But additionally, inference requires that this integration occur in a novel situation, not one that has been encountered before (see section "Practical Applications of Inferential Communication" for an example of how to study this form of inference). Furthermore, our definition of inference affords inferential thinking to the signaler, which allows a greater depth of cognitive engagement, including intended meaning from the signaler.

Second, there is no question that contextual pragmatics play a crucial role in the inferences recipients make in communicative exchanges. For instance, baboons may use the time of the day, the location, the activity or even the reproductive state of their groupmates to derive meaning from signals (Fischer, 2013; Fischer and Price, 2017). In our model of inferential communication, however, we open the door to mental state attribution and even the notion of common ground. We do not ascribe the most elaborate forms of mental state attribution and common ground to the inferences made in our model but propose that more basic levels of mental state attribution, such as knowledge state and past shared experiences, may be taken into account by both parties. This constitutes at least an entry point into a dimension that escapes contextual pragmatics, thus potentially bringing communicative exchanges closer to linguistic pragmatics. Note that our goal is not to downplay the importance of context in deriving meaning. On the contrary, contextual pragmatics play a fundamental role in the communicative exchanges of humans and primates (and possibly other animals), but we argue that there might be more to inference within animal communication than just contextual pragmatics, at least in certain interactions.

We also differentiate our model of inferential communication from the models of ostensive communication (Scott-Phillips, 2015; Moore, 2017; Heintz and Scott-Phillips, 2022), particularly with respect to the nature of inference and the depth of mental state attribution. Models of ostensive communication highlight the importance of inference in communicative exchanges, but they use inference in a much broader sense than we do in our model. Ostensive models also emphasize the role of complex mental state attribution, often articulated as informative and communicative intentions. We discuss and contrast these models with our own proposal in greater detail in Section "Beyond Inferential Communication: Ostensive Communication." For now suffice to say that we conceive inferential communication as the vital missing link between models of intentional and ostensive communication.

One of the main virtues of intentional communication is that it places flexibility and individual use of signals center-stage. However, the flexibility afforded by this model is rather limited. The origin of signals in intentional communication is either phylogenetic or ontogenetic ritualization. Phylogenetic ritualization produces species-specific signals potentially shared by all members of a species (and other closely related species). Signals *per se* are rather fixed, although their usage can show some flexibility, particularly in the gestural domain, in terms of when individuals choose to produce them, and whether they repeat them or replace with other signals in their repertoire when they fail (Liebal et al., 2014; Tomasello and Call, 2019 for review). This certainly shows some voluntary control over signals, but phylogenetic ritualization cannot produce new signals within an individual's lifetime. This is mainly the task of ontogenetic ritualization whereby two individual shape each other's behavior over repeated interactions so that they transform instrumental into communicative actions (Pika et al., 2005).

The production of novel signals is an important achievement, but ontogenetic ritualization is a slow process likely governed by associative learning. This means that new signals invariably require repeated interactions before they become fully functional. Attempts to document other forms of learning, most notably imitative learning, have failed to produce convincing evidence this form of learning is responsible for gesture acquisition in chimpanzees (Tomasello et al., 1997; Tennie et al., 2012). Inferential processes offer an alternative to ontogenetic ritualization and associative learning so that individuals can spontaneously invent gestures that others might be able to comprehend. Inference has been documented in numerous studies of physical cognition in primates (e.g., Hill et al., 2011; Petit et al., 2015; Völter and Call, 2017). Whether primates can also use inference in communicative situations is unclear but worth investigating. **Table 3** presents the types of inference that could be involved in primate communication. Each of these types requires increasing levels of cognitive sophistication. In the subsequent sections, we develop our proposal for inferential communication starting with situations involving social inferences in the absence of communication.

Social Inference

Of all the cognitive skills included in the model of inferential communication, the capacity for inference is both the most obvious and the most critical. Inferential communication is a system which demands a certain flexibility in interpretation of social interactions, where individuals must make leaps of understanding regarding the social behavior of another actor. One might call this "social inference," defined here as a situational understanding of another's actions beyond the available semantic information. Not restricted to communication, this ability includes successful interpretation of another's goals, intentions, or desires, in both cooperative and competitive contexts. Although social inference is not necessarily within the realm of communication, it is a vital prerequisite to inferential interpretation of another's communicative behavior. Social inference asks, "What does she want to do?" an open-ended question that relies on behavior, context, and inference in order to successfully attribute the ultimate goal to a set of actions performed by another.

Take, for example, our grooming primates. Now, rather than describing a communicative exchange, we can use their behavior to illustrate social inference. In this situation, Cindy grooms herself, producing species-typical grooming behaviors, such as plucking and licking certain areas of the body. She does not specifically intend to produce any particular signal, but she is observed by Louis, who makes inferences about her goals. Louis, observing her plucking behavior, could mentally represent her goal, which might be to alleviate an itch, clean a wound, or even to self-soothe after a tense encounter. Louis' inferential interpretation could be based on contextual cues (e.g., a visible wound, having witnessed a fight between Cindy and another individual, etc.), and/or past experience (Louis has groomed himself in the past and is aware of the benefits). Louis' capacity for inference, demonstrated here in his differential interpretation of Cindy's actions based on context, invokes the cognitive skill of goal attribution, which is not a requirement for the recipient in any of the previous models of communication. Additionally, Louis shows evidence of addressee awareness, in the sense that he is aware that he is not being addressed, which invites a different interpretation of Cindy's goal than if the behavior had been communicative and directed at him.

There is ample evidence for social inference in primates, including rational imitation, where great apes were less likely than human children to perform extraneous actions to complete a task, even when those actions had been demonstrated by a human actor (Call and Tomasello, 1998; Buttelmann et al., 2007, 2008). The apes appeared to infer the ultimate goal of the experimenter's actions and were able to produce a different, streamlined set of actions toward the same goal, rather than copying the experimenter's exact movements, indicating that they were able to use the experimenter's behavior to form a representation of their intentions. While perhaps reflecting less of an inclination toward social learning than human children, who readily imitated both the necessary and extraneous actions of the experimenter, these studies demonstrated that apes were able to infer the ultimate goal of the human's task, and thus eliminate unnecessary steps, suggesting a successful leap in understanding regarding the human's ultimate intention. Great apes also flexibly interpreted an experimenter's behavior in differing contexts, although the experimenter's actions were identical in both situations. Subjects were more likely to select one of two boxes when the experimenter "intentionally" dropped a marker on it versus when they "accidentally" dropped a marker on it (Call and Tomasello, 1998), which required inferences about the experimenter's goal when dropping the marker. Similarly, apes differentially adjusted their waiting behavior when experimenters were performing necessary actions on a puzzle box to retrieve food, compared with contexts where those same actions were superfluous, suggesting that they made inferences about the goal of those actions based on different contexts (Buttelmann et al., 2012). In each of these examples, despite identical semantic information, apes flexibly adjusted their responses (e.g., selection behaviors, waiting behaviors, begging behaviors) in response to different perceived goals from the experimenter. This evidence suggests that great apes have the ability to make pragmatic inferences about social behavior based on clues from context alone.

Evidence of social inference in primates is not limited to the interpretation side of social interactions. When it comes to production, both apes and monkeys show flexible adjustment of vocal signals based on the identity of the recipient (Cheney and Seyfarth, 2018). For example, chimpanzees produce food grunts toward "friends" more often than "non-friends" (Schel et al., 2013), and female baboons have been shown to selectively produce conciliatory grunts, mediated by the likelihood that the recipient will view their behavior as affiliative, where immediate past experience and long-term dominance dynamics appear to be the moderating factor (Cheney and Seyfarth, 2008). Audience effects such as these have been noted as evidence in reviews of intentional communication in primates (Liebal et al., 2006; Byrne et al., 2017), but they also present a potential case for inferential cognition, if and when these signals are voluntarily produced or withheld. While changes to vocal signals according to varying situational context (Seyfarth and Cheney, 2010 for review) are not enough to suggest social inference, variation in communication behavior regulated by *social* context, combined with voluntary control of these signals, allows that signalers may have an awareness that the intended message may be received differently by different individuals, depending on the existing social relationship with the specific partner. The possibility that primates can flexibly adjust communication behavior based on the varying potential outcomes from different recipients suggests that they have may be able to base these decisions on inferences from past social experiences, which goes beyond the realm of mere intentionality.

TABLE 3 | Social inference (non-communicative) and three types of inferential communication presented in ascending order of complexity in terms of the signal production and comprehension.

	Concept	Signaler's intention	Recipient's inference	Cognitive skills
Social Inference	Instrumental Action	I want to do X	What does she want to do?	Goal attribution
	Ambiguous Signal			Prosociality Informative Intention
Inferential Communication	Re-purposed Signal	I want **her** to do X (= $\mathbf{x_1 + x_2 + x_3}$) **for me**	What does she want **me** to do?	Innovation Context Rationalization
	New Signal			Iconicity Pantomime

Also depicted is the signaler's intention and recipient's understanding of those signals in reference to the intention communicated by the signaler.
Bold lettering represents the social and goal-directed nature of the signaler's intention and the recipient's inference.

Inferences Using Communicative Signals

Ambiguous Signals

In the case of fixed, semantic, unambiguous signals, advanced cognitive mechanisms are not necessarily required. In the case of flexible, ambiguous signals – those which are used in multiple contexts to mean different things – inference is a necessary component of interpretation. In order to apply social inference to the realm of communication, we must first consider the mentalized question at hand, for both the signaler and the recipient. The signaler asks, *"What do I want **him to do**?"* This question involves both an informative intention (that which she wants him to do), and a prosocial desire (the fact that she wants or needs him to do it, at no immediate benefit to himself). The recipient, on the other hand, asks, *"What does she want* ***me*** *to do?"* This question has an inherently prosocial component, which denotes a new cognitive skill, not explicitly required by past models, on the part of the recipient.

This arrangement requires the signaler to transmit a message which relies on the recipient's capacity for inference in order to be fully interpreted. She must establish attentional contact with the recipient to open a channel of communication between them, and must request help, providing instructions as to the nature of the help required. If she uses an ambiguous signal to provide instructions, however, the recipient must use inference to successfully interpret the message. For example, returning to Cindy and Louis, we now imagine a situation in which Cindy wishes to be groomed by Louis. She faces Louis, ensuring that he observes her, which opens the channel of communication. She taps her knuckles against the ground and bobs her body up and down, a gesture which is commonly used to initiate play, but has also been observed preceding grooming (Tomasello et al., 1997). In this gesture, she has both requested help from Louis, and provided instructions – she wants help in the form of grooming. She has, however, relied on her knowledge of the contextual inferences Louis is most likely to make, as well as the belief that Louis will be sufficiently motivated to help her. Louis, for his part, must attend to Cindy, be motivated to help, and understand the instructions, using context to disambiguate her gesture. If the communication is successful, then Louis will use contextual cues (for example, past experience with Cindy, in which they have rarely engaged in play) to interpret her instructions, and he will infer that she wants to be groomed (see **Figure 2**).

Regarding flexible interpretation of ambiguous behavior, there is evidence that primates are capable of such inferences. For example, great apes successfully differentiated between the same action from a human experimenter, producing more begging and impatience behaviors when the experimenter's action could be interpreted as "unwilling" to provide food versus "unable" to provide food (Call et al., 2004). From the production side, apes were able to modify the shape and location of their pointing behavior when their options were arranged such that a simple forward point would be ambiguous, suggesting that they recognized the need to disambiguate their pointing for the experimenter's successful comprehension and that they were able to apply that understanding to their actual gestures (Tauzin et al., 2020). Furthermore, great apes will monitor the success of a relatively ambiguous signal (e.g., begging), and elaborate with different, additional gestures (e.g., pointing at the desired option), if the desired outcome is not achieved (Leavens et al., 2005). This demonstrates a willingness to produce ambiguous signals, suggesting that the apes have some expectation that the signals will be successfully disambiguated by the experimenter, and also the capacity to choose whether or not to be more specific, at potentially higher cognitive cost to the signaler, if the ambiguous signal fails.

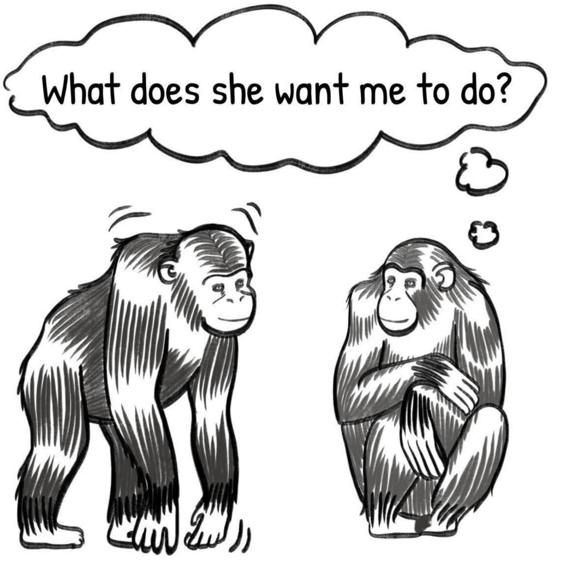


FIGURE 2 | Illustration of two primates engaged in a communicative exchange using ambiguous signals, depicting the recipient's inference under the inferential communication model. Illustration by Sadie Tenpas.

Re-purposed Signals

In situations where an ambiguous signal made unambiguous through inferential interpretation is not sufficient to thoroughly instruct the recipient, the signaler may turn to other resources to produce an instructive signal. One possible approach is to use an existing signal within the communicative repertoire, but in a brand-new context, relying on the inferential capability of the recipient to interpret the familiar signal in a new way. The situational question remains the same for both the signaler and the recipient – "What do I want him to do?/What does she want me to do?" – but new cognitive skills are required at this level of complexity. In addition to the required capacity for inference and prosocial behavior, the signaler and the recipient must both take a creative leap and rationalize the otherwise nonsensical use of the signal in the current situation, giving it new meaning.

If we follow primates Cindy and Louis into a new situation, an experimental setting in which they must work together to open a puzzle box, we can hypothesize an interaction using this form of inferential communication. Cindy wants Louis to help her open a locked puzzle box, which can be achieved by turning two wheels, simultaneously, at opposite ends of the box. As neither can reach both wheels, they must coordinate to solve this problem. As before, Cindy must establish a communicative channel with Louis, by looking at him and ensuring that he sees her. As no fixed, semantic gesture exists in their shared repertoire to communicate "help me open this box," Cindy produces a gesture more typically used to beg for food, a mouth stroke (Tomasello et al., 1997). The gesture is nonsensical in this context, as Louis has no food to offer her. Instead, Cindy has engaged in creative use of this gesture to encourage Louis to open the box with her. If this exchange were to be successful, Louis would correctly infer that Cindy does not want to share food, rather, he would rationalize the otherwise pointless gesture to a new meaning, and if sufficiently motivated, help Cindy open the box (see **Figure 3**).

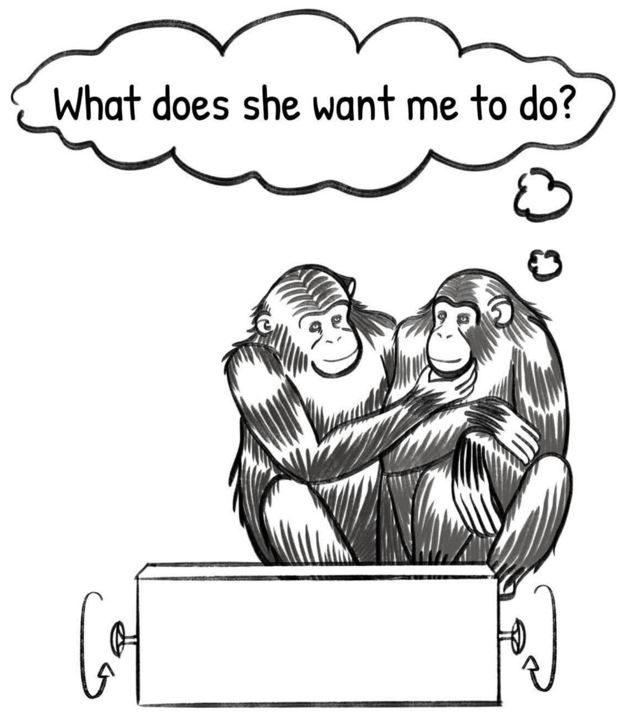


FIGURE 3 | Illustration of two primates engaged in a communicative exchange using re-purposed signals, depicting the recipient's inference under the inferential communication model. Illustration by Sadie Tenpas.

This exchange relies on creative re-purposing of existing gestures, a capacity which has not been conclusively demonstrated in primates, but which can be hypothetically proposed in the model of inferential communication. Armed with this framework, it is possible to design experiments which more specifically demand this ability, to explore the cognitive skill and its presence or absence in primates.

New Signals

In its final possible level of cognitive complexity, inferential communication provides a platform for two actors to create an entirely new signal, rationalized and understood by both purely based on the context and their own capacity for inference. Still adhering to the question, "What do I want him to do?" the signaler instructs the recipient using an iconic gesture – one that does not exist in the known repertoire of the individual, and which pantomimes the action she is requesting that the recipient perform. This iteration of inferential communication adds two specific cognitive skills not required for earlier levels: iconicity and pantomime, which are necessary for both the signaler and recipient.

If we return to Cindy, Louis, and the puzzle box, we can imagine a situation in which Cindy establishes that she has Louis' attention, and then turns her hand in the air, miming the turning of the wheels on the puzzle box. Louis, observing this pantomime, interprets the gesture as an iconic representation of the desired action, understands Cindy's request for help and the instructions she has given, and helps her open the box (see **Figure 4**).

While this type of interaction has not been systematically documented in primates, and it is unlikely that this type of interaction is common, preferred, or cognitively efficient for non-human animals, the question remains as to whether primates could exhibit these cognitive abilities if there were no other way to solve the problem. There is some anecdotal evidence that primates are capable of the two new cognitive skills seen here – iconicity and pantomime. Grosse et al. (2015) found that one chimpanzee, who had been partially reared by humans, engaged in an iconic gesture when a human experimenter required instruction to operate an apparatus. Additionally, great apes have been anecdotally observed engaging in pretend play, with or without the assistance of objects, suggesting some possibility of iconic representation of objects (Gómez, 2005). On the comprehension side of these abilities, great apes have been shown to learn locations associated with iconic gestures faster than locations associated with arbitrary gestures, suggesting that they have some ability to link the iconic nature of those gestures to their representational meaning (Bohn et al., 2016a). It is noteworthy that production of iconic signals, iconic play, and imitation of pantomimed gestures can be scaffolded with the support of physical objects, removing the requirement of intransitivity that is intrinsic to true pantomime (e.g., Call, 2001; Gómez, 2005; Tennie et al., 2012). This suggests that iconicity and pantomime are challenging cognitive skills for great apes and would require substantial prosocial motivation or necessity in order to be a cognitively efficient mechanism. Whether these

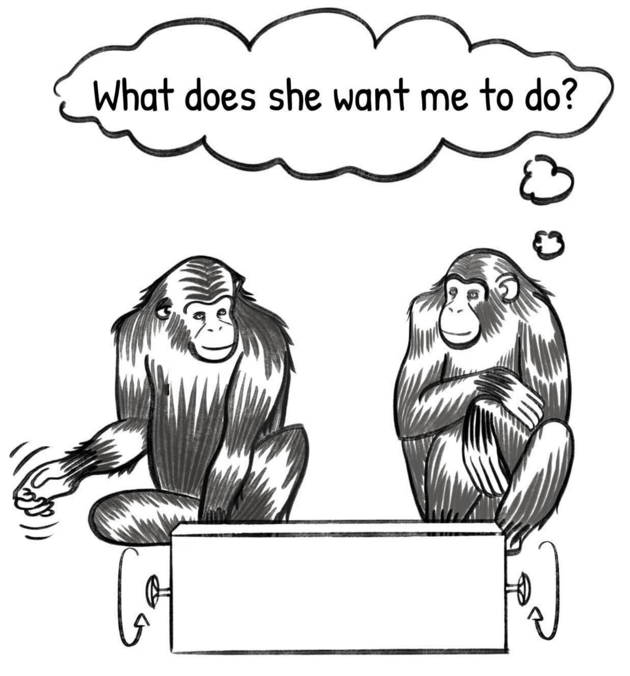


FIGURE 4 | Illustration of two primates engaged in a communicative exchange using newly invented signals, depicting the recipient's inference under the inferential communication model. Illustration by Sadie Tenpas.

anecdotal cases could be transformed into robust evidence of these cognitive skills in primates, especially in the absence of scaffolding, is unknown, but the question itself presents an exciting example of the investigations and experiments that become possible on the tails of inferential communication. It also invites the question of prosociality in primates, given that prosocial action is a critical component of the success of any inferential communication exchange, but especially those with increasingly difficult cognitive demands.

Prosocial Motivation

One of the basic tenets of human communication is that it is a prosocial and cooperative enterprise (Hare, 2017). Although prosocial motivation can foster communicative exchanges, and it may be needed for language acquisition, we argue that it is not necessary for inferential communication because effective exchanges can occur even in the absence truly prosocial (i.e., altruistic) behavior. Apparently prosocial actions, required of both actors in inferential communication exchanges, can be understood with several different motivations in mind. On the surface, prosociality is defined as any action, whether requested or performed, that one actor completes for the benefit of another, with either no benefit or actual cost to themselves (Cronin, 2012). In practice, however, apparently prosocial actions, those performed at cost of one individual and benefit of another, may have motivations other than pure altruism. An individual could be motivated to behave in a prosocial manner due to a cost-benefit analysis, in which the continuing annoyance or harassment from the requesting individual is a greater cost than performing the action, and the actor is therefore sufficiently prosocially, if not altruistically, motivated. Alternatively, the actor could incur hidden ultimate benefits, such as augmented reputation, hopes for reciprocation, or, in humans, a proximate benefit of internal moral good feeling, which offset the apparent cost of the prosocial action. Thus, when we argue that prosociality is a requirement for successful inferential communication, we refer to the broad spectrum of motivations that could lead to apparently prosocial action. Apparently prosocial actions, whether altruistic or otherwise, have been observed in several species of primate, from tamarins (Cronin et al., 2010) to macaques (Massen et al., 2010), and to great apes (Pelé et al., 2009).

Altruistic prosocial motivation and willingness to engage in the inherently cooperative act of communication become more imperative as the cognitive load of the exchange increases. For the both the signaler and the recipient in a communicative exchange, the cost-benefit analysis of the effort to produce or interpret a communicative signal changes depending on the difficulty associated with interpreting the message. For fixed signals and ambiguous but commonly used signals, the cognitive effort may not override the beneficial outcome of the signaler and the potential hidden benefits for the recipient. When the more demanding cognitive skills mentioned above – creativity, rationalizing re-purposed signals, iconicity, and pantomime – are necessary for the exchange, the outcome must be more beneficial for both parties. Particularly for the recipient, it may be that this type of cognitive load is only worth the effort if the individual is truly altruistically motivated, a phenomenon which remains debated in primates (Cronin, 2012). Perhaps prosocial motivation is crucial to decode opaque messages that otherwise are simply not understood and consequently, ignored. It is possible, therefore, that lack of existing evidence for these later levels of inferential communication is caused by twofold limitations. First, the necessity for truly altruistic motivations, which appear to occur sparsely, if at all, in primates, and second, the difficulty of the cognitive mechanisms at play. It is possible, however, that in the presence of sufficient prosocial motivation, primates could produce and interpret these types of cognitively complex signals.

BEYOND INFERENTIAL COMMUNICATION: OSTENSIVE COMMUNICATION

Although our focus is on inferential communication, it is critical to discuss ostensive communication for the sake of contrast and completion. Ostensive communication makes the leap from social inferences to communicative inferences – specifically, inferences about communicative intention (see **Table 2**). Communicative intention is traditionally understood as a mental state function, in which the communicator not only knows the mental state of the recipient, but consciously intends to manipulate that mental state by making their own informative intention manifest. This is combined with the recipient's recognition that the communicator has an informative intention, which prompts the recipient *via* the presumption of relevance to make inferences about the meaning of the message based on contextual cues and mental states (Scott-Phillips, 2015). The capacity of primates to produce and comprehend communicative intention under this mentalistic definition is not clearly understood – it has yet to be conclusively observed or experimentally demonstrated in great apes, and it is seldom investigated in monkeys and prosimians (Moore, 2016). Some researchers assert that this cognitive capacity is unlikely to exist in primates, given the sufficiency of a sparser, more goal-directed and intentional model of communication to explain most communicative exchanges primates (Fischer and Price, 2017), and given that primates frequently fail tasks which require production or comprehension of communicative intention (Tomasello, 2008). This mentalistic definition of communicative intention requires recursive mental state attribution, including fourth-order theory of mind (Scott-Phillips, 2014), which many regard as too complex for primates.

Ostensive communication asks the question, "*What does she want* ***to tell me*** *to do?*" where not only the informative intention, but also the communicative intention, is manifest to the recipient. In our hypothetical primate example, Cindy wishes to be groomed by Louis. In order to accomplish this, Cindy makes inferences about Louis' current mental state – his willingness to groom her, his awareness that she wants to be groomed, and their existing shared knowledge – and knowingly and intentionally sets out to alter his mental state with her message, such that he becomes aware that she wants to be groomed and is motivated to do so. Louis, likewise, as the recipient, must be aware that Cindy is attempting to alter his mental state, and uses that knowledge to make inferences about what she wants, based on the situation context.

Our model of inferential communication shares two key aspects with the model of ostensive communication: inference and mental state attribution. However, it differs in the type of inferences that it uses, and it lacks the most complex forms of mental attribution, particularly communicative intentions. Models of ostensive communication use inference with two meanings, one very broad (all communication involves some form of inference) and one rather narrow. For instance, Scott-Phillips (2015) describes inference as the interpretation of meaning based on evidence of informative and communicative intentions of the signaler. In our model of inferential communication, we do not ascribe expression or interpretation of communicative intention to either actor, but rather suggest that the signaler is relying on the recipient to make inferences about their goals (i.e., informative intention), rather than their communicative act itself. Thus, although we agree with Scott-Phillips (2015), Heintz and Scott-Phillips (2022) that communicative intentions may be beyond the capabilities of non-human animals, we argue that informative intentions might not be – signalers can express their goals informatively, but do not make their intentions manifest and recipients do not necessarily use presumption of relevance to infer meaning.

Recently, Heintz and Scott-Phillips (2022) distinguished between ‘intentional expression’ defined as the expression of mental states (e.g., a signaler may indicate what she wants to a recipient) and ostensive communication that requires making informative intentions manifest. We view intentional expression as similar (if not equivalent) to what we are calling inferential communication, except that we argue that informative intentions (perhaps in a more rudimentary form) are conveyed by signalers, but communicative intentions are not. Other authors have argued that non-human animals may even exhibit some forms of communicative intention. Moore (2017) argues that primates may indeed exhibit a form of Gricean, ostensive-inferential communication, but emphasizes the role of awareness of address on the part of the recipient to provide the context for interpretation, which is followed by inferences about the goal of the signaler.

The complexity of mental state attribution required by full-blown (human) ostensive communication is not yet evidenced in primates. Like the more complex levels of inferential communication, it is possible that both the cognitive and the prosocial demands are too great for the majority of communicative exchanges between primates. Perhaps, with evidence taken from an inferential communication framework, it might be possible, in the near future, to design experiments which better establish the limits of primate mental state attribution, to further bridge the gap between language-oriented developmental literature, with rich, Gricean interpretations of communication, and comparative literature, where interpretations are currently limited to description and suggestion of cognitive engagement. The model of inferential communication, when applied to observations and experiments in non-human animal behavior, presents the opportunity to ask theoretical questions about flexible communicative problem solving, theory of mind, and communicative intention.

PRACTICAL APPLICATIONS OF INFERENTIAL COMMUNICATION

Any newly proposed theoretical framework to study animal communication, has to consider its practical applications: what can this model offer, how can it be examined empirically, and which species are more likely to display it? In this section, we highlight some of the specific applications of inferential communication and propose some examples of experimental designs to test whether recipients, but also signalers, use inference in their communicative exchanges. To do so, we present three tasks, one using vocal communication and focused on inferred meaning, and the two others using gestures and focused on intended meaning. We intentionally provide methods examining both vocal and gestural communication in order to illustrate the complementary roles played by vocal and gestural communication in elucidating the intended and inferred meaning of ambiguous, re-purposed and novel signals. We close this section by outlining our criteria for determining whether a species might be a good candidate to investigate inferential communication.

A key application of inferential communication is to explain the origins of some gestures. It is recognized that gestures can arise *via* phylogenetic or ontogenetic ritualization (Cartmill and Hobaiter, 2019). Another mainly discarded form is third-person imitation (Tomasello et al., 1997; Tennie et al., 2009), although language trained apes have been reported to learn some signs by imitation (Fouts, 1972; Gardner et al., 1989). We propose that inference could serve as a fourth form of gesture acquisition; creating a new gesture to indicate old or new meaning, or less demanding, re-purposing a gesture, and here the work is in using it with a different meaning and especially interpreting it. Different from phylogenetic ritualization, where a successful gesture is preserved and inherited in the innate repertoire, and ontogenetic ritualization, where a gesture develops from repeated use of action-oriented movements, inferential development could explain gestures which originate as iconic or re-purposed movements and are practiced and used until they are semantically established between two or more individuals. Notably, this understanding of the origin of gestures would differentiate between ritualized gestures which iconically evoke the requested action, but evolve from the actual occurrence of the action, and inferred gestures, which originate from an iconic, pantomimed representation of the action.

It is crucial that inferential communication, as an origin of gestures and as a cognitive process, be explored experimentally. Novelty is an essential part of the development of new gestures; if two individuals use a gesture repeatedly, there is no need to invoke inference. Inference need not necessarily be applied in all communicative exchanges, but when the system is perturbed (new conditions or old conditions no longer apply) it can play a crucial role in the success of communication. Once invented by inference, a gesture may be used repeatedly, which can quickly mask its origins. Thus, experiments are critical in order to observe the emergence of new forms of communication.

In designing experiments to test inferential communication, it is essential to ensure that the task requires true inference – the integration of known information to understand a new scenario. Likewise, the experiment should require the use of pragmatic information on the part of both parties, not just on the order of situational context clues, but on the mentalized level of knowledge state, private interaction history, or individual preferences. Our first proposed experiment applies these two criteria to interpreting ambiguous vocal signals. We imagine an experimental setting in which the recipient of the communication knows two established pieces of information, which may have been learned by past inference, association, or simply occurred as a result of the individual's maturation. The point is that the origin of the two pieces of information is not so relevant in our example. First, the recipient, a primate in this example, must be familiar with the species-specific vocalizations that individuals produce when they discover a cache of highly preferred food. Second, the recipient must be familiar with the individual food preferences of a particular groupmate. For instance, Cindy, our recipient, knows that Louis likes bananas but does not like grapes. This is something that she could have learned by observing Louis' feeding patterns: always eating bananas with gusto but ignoring grapes, when both are available, and even when bananas are gone, Louis shows no interest in grapes still available. Cindy is also familiar with their species-specific food calls, which are associated not with a particular type of food, but with the discovery, prior to, but not during eating, of a cache of food.

In the test condition, Cindy is shown that one of two foods is hidden behind a bush, but she is not shown which type of food. Louis then appears and produces a food call upon encountering the food cache. If Cindy is indeed capable of integrating multiple pieces of known information to infer meaning in this new situation, we predict that she should infer that there are bananas behind the bush. Moreover, she should be surprised, in this instance, to search and discover grapes behind the bush, and this response pattern should be reversed if the caller was an individual who likes grapes and does not like bananas. Notably, experiments like this allows examination of the first exposure to this novel situation, which is important for evaluating inference. Associative processes require at least one event for learning to occur, which means the recipient's reaction on initial exposure in the proposed experiment is a measure of true inference. As far as we know, this proposed experiment has not been done (but see Shorland, 2018 for a similar experimental paradigm), but we already know that chimpanzees integrate the food preferences of others and their visual access when choosing between two experiments – selecting the one which will give them the most favorable outcome (Eckert et al., 2018). This experiment would test whether they could extend this ability to integrate information to inference in communicative exchanges.

Compared to the work investigating recipient comprehension, much less has been done examining the inferential abilities of signalers, with some authors arguing that primate signalers do not intend meaning, recipients just infer it (e.g., Fischer, 2013; Fischer and Price, 2017). This is a sensible proposition given that primate vocal signals are fixed, apart from variation in the timing and context of their use. Such inflexibility in vocal production may not permit primate signalers much opportunity to imbue meaning to their signals. Gestures, on the other hand, are quite different in terms of their production. Gestures are grounded in bodily action; they are much more flexible than vocalizations, which opens the door for flexible variation that changes the potential interpretation of the signal. This flexibility also permits the creation of novel signals or the re-purposing of old signals to a novel use. Therefore, we challenge the idea that primate signalers in general do not ever intend meaning and argue that this conclusion may have resulted from asking this question from the perspective of vocal communication only.

The literature already contains some studies illustrating this point – examples we would argue indicate that signalers communicate intended meaning. For instance, Bohn et al. (2016b) found that great apes used a pointing gesture in an unusual way (pointing to an empty dish) to request food that was no longer in that dish. Pointing to an empty dish is atypical for apes, especially given that another dish containing food that was less preferred, but otherwise perfectly acceptable (they always ate this food in control trials), was present. Special care was taken to avoid training the apes to point to an empty container in the pre-test, where they witnessed that the experimenter got up as soon as the food was depleted, left the room, and brought in more food, without giving the subject a chance to point to the empty container. Importantly, apes only used this unusual gesture with an experimenter who had brought food in the past as soon as food had been depleted but not with an experimenter who had given them food but not brought it in the first place.

When we analyze the key features of this case, we conclude that pointing to an empty container qualifies as re-purposing a familiar gesture to communicate about an absent referent. First, pointing with extended fingers, unlike vocalizations, is not a species-specific gesture, but one that is acquired in contact with humans, thus showing some degree of flexibility in gestural acquisition. Second, the pointing gesture is directed at referents (e.g., food item) that are present (even when they are hidden), not to empty dishes, which suggests that the apes in the experiment were using the gesture in a novel way. Third, apes used the pointing to the empty container only with the experimenter that they had experienced bringing food, and not with others, suggesting that they accounted for the private interaction history between themselves and the experimenters in order to inform their knowledge of whether the gesture to the absent entity would be meaningful. There are other examples in which apes communicated what they wanted by re-purposing an action to request help from an experimenter. The bonobo Kanzi pounded on a nut to request that an experimenter to crack it open (Savage-Rumbaugh et al., 1986). In anecdotal observations, juvenile gorillas physically guided human researchers toward locked doors, using gaze-alternation throughout the movements, presumably to indicate to human observers what needed to be done (Gómez, 1990). The fact that in these instances apes established eye contact with the human experimenter when performing their actions toward the door and reduced their rate of these door-approaching actions when the experimenter left the room strongly suggests that the apes were using those acts to communicate with the experimenter, and not purely as a goal-oriented mechanism.

Intended meaning could also theoretically occur in more complex forms of communication, whereby apes invent a new gesture by, for example, pantomiming an action to indicate the tool that they require to obtain food. Yamamoto et al. (2012) reported that chimpanzees transferred tools to their partners following requests. Signalers used a hand begging gesture and recipients, who could see the kind of tasks that signalers were facing, selected the correct tool from an assortment of various tools and gave them to the signaler. When the recipient's view of the signaler's task was blocked, however, they handed tools randomly. This means that the begging gesture itself did not carry meaning about the type of tool. Context provided that information because the recipient could see the tool that was needed. Thus, the burden of decoding the message fell on the recipient who used contextual cues (the type of apparatus present) to infer meaning. This level of inference is based on contextual pragmatics, not mental states, but it begs the question: would the signaler become more specific in her request, and perhaps even invent a novel gesture by pantomiming the use of a specific tool, if the lack of contextual information persisted over time? We think that this might be asking too much from signalers, who seem to have trouble producing intransitive actions in imitation studies (Tennie et al., 2012). Thus, a pounding action to indicate a stone hammer might be outside of the spontaneous repertoire of primates, but if the intransitive action could be scaffolded with transitive elements, it might be possible that primates could gesture with intended meaning using novel signals.

In this potential experimental arrangement, with the possibility of scaffolded novel gestures, it is possible to examine whether signalers would take the context into account when producing their signals, which would suggest an awareness of the inferences they can reasonably expect the recipient to make. If the context already provides enough information about their intended meaning, would their signals become less specific, particularly when more specific signals are costlier to produce? Conversely, if contextual cues are ambiguous, would signals become more specific? There exists some experimental evidence that apes use pointing variations to disambiguate between two food items when the higher-value food was placed behind a lower value food and subjects were asked to select their preferred food, *via* pointing (Tauzin et al., 2020). A similar paradigm could investigate whether apes use modified pointing gestures to disambiguate between choices where the context is identical, but their knowledge of the recipients' past experiences or preferences is varied. For example, if one experimenter is known to always provide the higher-value food regardless of the spatial arrangement of the plates, but another experimenter is new to the situation and has no expected pattern of behavior, will the subject use modified pointing gestures to disambiguate their choice with the new experimenter, but not the familiar one?

We now turn our attention to criteria for determining which taxa and which species might be more likely to display inferential communication. Based on the examples that we have given, primates, and particularly great apes seem suitable candidates to investigate the existence of inferential communication. While the communicative behavior of many species is sufficiently captured by explanations found in the foundational models of communication, it is possible that other species, apart from primates, may also be capable of inferential communication provided they possess the required cognitive prerequisites. We propose three such prerequisites: goal-directed communication, general inferential reasoning abilities, and non-communicative social inference. If evidence of these abilities is found in any species, regardless of taxa, it is possible that inferential communication may be within their capacity as well. For example, there is a body of evidence that canines exhibit intentional communication (e.g., Rossi and Ades, 2008) and social inference (e.g., Bräuer et al., 2006). African gray parrots have been shown to exhibit general inferential abilities (e.g., Schloegl et al., 2012; Pepperberg et al., 2013, 2018), and there is some evidence to suggest that they possess the capacity for intentional communication as well (e.g., Pepperberg, 2004). These groups may, therefore, be promising candidates for inferential communication, but rigorous testing of the above pre-requisites would be necessary before investigations of inferential communication could be practically conducted in any of these groups. We do not suggest that any species meeting these criteria is *de facto* likely to use inferential communication, we merely suggest that possession of these prerequisites may serve to determine whether that species is worth closer investigation.

In sum, we have proposed several ways by which inferential communication can be used to investigate inferred meaning by recipients using true inference – the integration of information to be applied to a new scenario – as well as contextual clues based not only on situational factors, but also on the mental state of the signaler. Furthermore, we have highlighted some tasks already present in the literature that we believe test for intended meaning on the part of a signaler, and proposed ways that they could be modified to new tasks to investigate whether primates can integrate simple theory of mind into their accounting of context, and whether signalers can account for such context while producing more complex (in terms of iconicity) forms of communication. We also indicated that a species possessing goal-directed communication, general inferential reasoning abilities, and/or non-communicative social inference (with all three abilities constituting the strongest foundation) would be a good starting point to investigate inferential communication.

CONCLUDING REMARKS

The field of animal communication has made considerable progress since the appearance of the early ethological models purely based on behavior. Much of this progress has occurred as a consequence of the development of cognitive models of animal communication. In what has otherwise been a progressive increase in cognitive sophistication aimed at explaining flexible communication, we think that the field now runs a risk of stagnation due to the rejection of any form of mental state attribution in communication (Townsend et al., 2017). In this paper we have argued that we need a more thorough and detailed understanding of mentalizing in communication, particularly for species that are flexible communicators, and especially when those data are subsequently used to make inferences about the evolution of language. Without mentalizing in models of animal communication, the gap between animal and human communication might be too wide to bridge.

We submit the model of inferential communication as a way forward – a way to progress from descriptions of potential cognitive outcomes to considerations of the actual cognitive mechanisms driving them. Evidence of cognitive forms of communication in primates, especially intentional and referential exchanges, combined with evidence of social inference such as goal attribution, leads us to propose that primates (and perhaps other species too) may have the capacity to make inferences within communicative exchanges. The idea of inference playing a role in animal communication is not new, but we argue that its potential importance and scope has not been fully realized because inference has often been conflated with other mechanisms. Moreover, we propose that investigation of inferences involving the integration of disparate pieces of information, some not based on contextual cues, may provide new insights into the mechanisms underlying the complexity and flexibility of primate communication. Our model invites a rich interpretation of the cognitive mechanisms surrounding communication by challenging the idea that meaning is drawn exclusively from a set of rules or semantics, or from conditional discrimination between situations, which might otherwise suggest simplistic associative learning or hardwired signal-response connections. We also decouple informative intention from communicative intention and suggest that it is possible for actors in a communicative exchange to engage with simple mental state attribution and expression of goals, absent the recursive levels of theory of mind found in ostensive models of communication.

The model of inferential communication is a multi-level framework, beginning with social inferences regarding non-communicative behavior and extending to communicative inferences regarding how signals are used and interpreted, including consideration of the motivation underlying communicative exchanges. With regard to signals, we have illustrated the inferential approach to interpreting ambiguous signals, re-purposing old signals, and creating new ones. Each level shares the fundamental requirement that both individuals, the signaler and the recipient, must make leaps of interpretation for successful communication. For some of these levels, there is already some evidence suggesting that primates might be capable of communicative inferences, but for other levels there is only anecdotal or even negative evidence. Furthermore we have proposed ways in which these ideas could be tested using new tasks or by modifying existing ones. With regard to motivation we have argued that a prosocial motivation is not strictly necessary for this form of communication to arise because it can hijack other motivational systems for the same successful outcome, but if present, it may facilitate successful communication involving the production of novel and initially opaque signals.

Our proposal extends beyond current approaches to referential and intentional communication but stops short of ostensive communication. Although we do not rule out *a priori* the possibility that ostensive communication could occur in primates, we suggest that before tackling this issue, is important to explore the possibility of inferential communication, which is in some ways a pre-requisite for ostensive communication. Our proposal therefore does not qualify as mentalistic communication in the Gricean sense (Grice, 1957, 1969, 1989) but unlike Townsend et al. (2017) it does not flatly reject the importance of some forms of mentalizing, which we incorporate to our model. Namely, we argue that goal attribution, visual perspective taking, and knowledge attribution may play an important role in the inferences that individuals make in their communicative exchanges.

Finally, our endorsement of inferential communication should not be taken as an indication that we believe primates engage in inferential communication in every communicative exchange. Instead, we propose that individuals mainly engage inferential communication when routine conditions change, and new solutions are required. In this sense, engaging inferential communication is analogous to engaging cognitive control and monitoring mechanisms in problem solving following the perturbation of a previously stable system. We believe that inferential communication is ideally placed to bridge the gap between the intentional and the ostensive model of communication, something that it is particularly important for those wishing to make inferences regarding the evolution of language. It is a framework that we hope will contribute to more precise descriptions of phenomena we have already witnessed in primates and promote new insights into the complexity of animal communication. It is a toolkit – a perspective that we hope will empower researchers to take a more productive approach to animal communication, both in design and interpretation.

AUTHOR CONTRIBUTIONS

Both authors contributed jointly to the conceptual and theoretical development of the model of communication presented in the manuscript, contributed to the writing and reviewing each draft of the manuscript, and approved the submitted version.

FUNDING

This work was supported by the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC Grant 609819 (SOMICS).

ACKNOWLEDGMENTS

We thank Sadie Tenpas for creating the illustrations displayed in this manuscript. We also thank the two reviewers, whose comments and suggestions helped improve and clarify the manuscript. We also thank Christophe Heintz and Julia Fischer for their valuable input on the ideas presented in this work.

REFERENCES

Arbib, M. A., Liebal, K., and Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Curr. Anthropol.* 49, 1053–1076. doi: 10.1086/593015

Bohn, M., Call, J., and Tomasello, M. (2016a). Comprehension of iconic gestures by chimpanzees and human children. *J. Exp. Child Psychol.* 142, 1–17. doi: 10.1016/j.jecp.2015.09.001

Bohn, M., Call, J., and Tomasello, M. (2016b). The role of past interactions in great apes' communication about absent entities. *J. Comp. Psychol.* 130, 351–357. doi: 10.1037/com0000042

Bräuer, J., Kaminski, J., Riedel, J., Call, J., and Tomasello, M. (2006). Making inferences about the location of hidden food: social dog, causal ape. *J. Comp. Psychol.* 120, 38–47. doi: 10.1037/0735-7036.120.1.38

Bretherton, I., and Bates, E. (1979). The emergence of intentional communication. *New Dir. Child Adolesc. Dev.* 1979, 81–100. doi: 10.1017/S0305000912000359

Buttelmann, D., Carpenter, M., Call, J., and Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Dev. Sci.* 10, F31–F38. doi: 10.1111/j.1467-7687.2007.00630.x

Buttelmann, D., Carpenter, M., Call, J., and Tomasello, M. (2008). Rational tool use and tool choice in human infants and great apes. *Child Dev.* 79, 609–626. doi: 10.1111/j.1467-8624.2008.01146.x

Buttelmann, D., Schütte, S., Carpenter, M., Call, J., and Tomasello, M. (2012). Great apes infer others' goals based on context. *Anim. Cogn.* 15, 1037–1053. doi: 10.1007/s10071-012-0528-4

Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., and Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Anim. Cogn.* 20, 755–769. doi: 10.1007/s10071-017-1096-4

Call, J. (2001). Body imitation in an enculturated orangutan (Pongo pygmaeus). *Cybern. Syst.* 32, 97–119. doi: 10.1080/019697201300001821

Call, J., Hare, B., Carpenter, M., and Tomasello, M. (2004). 'Unwilling' versus 'unable': chimpanzees' understanding of human intentional action. *Dev. Sci.* 7, 488–498. doi: 10.1111/j.1467-7687.2004.00368.x

Call, J., and Tomasello, M. (eds) (2007). *The Gestural Communication of Apes and Monkeys*. Oxfordshire: Taylor & Francis Group.

Call, J., and Tomasello, M. (1998). Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *J. Comp. Psychol.* 112, 192–206. doi: 10.1037/0735-7036.112.2.192

Cartmill, E. A., and Hobaiter, C. (2019). Developmental perspectives on primate gesture: 100 years in the making. *Anim. Cogn.* 22, 453–459. doi: 10.1007/s10071-019-01279-w

Cheney, D. L., and Seyfarth, R. M. (2008). *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago: University of Chicago Press.

Cheney, D. L., and Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1974–1979. doi: 10.1073/pnas.1717572115

Cronin, K. A. (2012). "Cognitive aspects of prosocial behavior in nonhuman primates," in *Encyclopedia of the Sciences of Learning*, ed. N. M. Seel (Boston: Springer), 581–583.

Cronin, K. A., Schroeder, K. K., and Snowdon, C. T. (2010). Prosocial behaviour emerges independent of reciprocity in cottontop tamarins. *Proc. R. Soc. B Biol. Sci.* 277, 3845–3851. doi: 10.1098/rspb.2010.0879

Dawkins, R., and Krebs, J. R. (1978). "Animal signals: information or manipulation," in *Behavioural Ecology: An Evolutionary Approach*, eds J. R. Krebs and N. B. Davies (Oxford: Blackwell), 282–309.

Eckert, J., Rakoczy, H., Call, J., Herrmann, E., and Hanus, D. (2018). Chimpanzees consider humans' psychological states when drawing statistical inferences. *Curr. Biol.* 28, 1959–1963. doi: 10.1016/j.cub.2018.04.077

Fischer, J. (2013). "Information, inference and meaning in primate vocal behaviour," in *Animal Communication Theory: Information and Influence*, ed. U. Stegmann (Cambridge: Cambridge University Press), 297–319.

Fischer, J., and Price, T. (2017). Meaning, intention, and inference in primate vocal communication. *Neurosci. Biobehav. Rev.* 82, 22–31. doi: 10.1016/j.neubiorev.2016.10.014

Fitch, W. T. (2015). Evolving pragmatics. *Curr. Biol.* 25, R1110–R1112.

Fouts, R. S. (1972). Use of guidance in teaching sign language to a chimpanzee (*Pan troglodytes*). *J. Comp. Physiol. Psychol.* 80, 515–522. doi: 10.1037/h0032989

Gardner, R. A., Gardner, B. T., and Van Cantfort, T. E. (eds) (1989). *Teaching Sign Language to Chimpanzees*. New York: Suny Press.

Gómez, J. C. (1990). "The emergence of intentional communication as a problem-solving strategy in the gorilla," in *Language" and Intelligence in Monkeys and Apes*, eds S. T. Parker and K. R. Gibson (Cambridge: Cambridge University Press), 333–355.

Gómez, J. C. (2005). "Joint Attention and the Notion of Subject: Insights from Apes, Normal Children, and Children with Autism," in *Joint Attention: Communication and Other Minds: Issues in Philosophy and Psychology*, eds N. Eilan, C. Hoerl, T. McCormack, and J. Roessler (Oxford: Oxford University Press), 65–84. doi: 10.1093/acprof:oso/9780199245635.003.0004

Graham, K. E., Wilke, C., Lahiff, N. J., and Slocombe, K. E. (2020). Scratching beneath the surface: intentionality in great ape signal production. *Philos. Trans. R. Soc. B* 375:20180403. doi: 10.1098/rstb.2018.0403

Grice, H. P. (1957). Meaning. *Philos. Rev.* 66, 377–388.

Grice, H. P. (1969). Utterer's meaning and intentions. *Philos. Rev.* 78, 147–177.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Grosse, K., Call, J., Carpenter, M., and Tomasello, M. (2015). Differences in the ability of apes and children to instruct others using gestures. *Lang. Learn. Dev.* 11, 310–330. doi: 10.1080/15475441.2014.955246

Hare, B. (2017). Survival of the friendliest: *Homo sapiens* evolved *via* selection for prosociality. *Annu. Rev. Psychol.* 68, 155–186. doi: 10.1146/annurev-psych-010416-044201

Heintz, C., and Scott-Phillips, T. (2022). Expression unleashed: the evolutionary & cognitive foundations of human communication. *Behav. Brain Sci.* 1–46.

Hewes, G. W., Andrew, R. J., Carini, L., Choe, H., Gardner, R. A., Kortlandt, A., et al. (1973). Primate communication and the gestural origin of language [and comments and reply]. *Curr. Anthropol.* 14, 5–24.

Hill, A., Collier-Baker, E., and Suddendorf, T. (2011). Inferential reasoning by exclusion in great apes, lesser apes, and spider monkeys. *J. Comp. Psychol.* 125, 91–103. doi: 10.1037/a0020867

Hobaiter, C., and Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Curr. Biol.* 24, 1596–1600. doi: 10.1016/j.cub.2014.05.066

Hopkins, W. D., Taglialatela, J. P., and Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Anim. Behav.* 73, 281–286. doi: 10.1016/j.anbehav.2006.08.004

Krebs, J. R., and Dawkins, R. (1984). "Animal Signals: Mind Reading and Manipulation," in *Behavioural Ecology: An Evolutionary Approach*, eds J. R. Krebs and N. B. Davies (Oxford: Blackwell).

Leavens, D. A., Hostetter, A. B., Wesley, M. J., and Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Anim. Behav.* 67, 467–476. doi: 10.1016/j.anbehav.2003.04.007

Leavens, D. A., Russell, J. L., and Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Dev.* 76, 291–306. doi: 10.1111/j.1467-8624.2005.00845.x

Liebal, K., Pika, S., and Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6, 1–38.

Liebal, K., Waller, B. M., Slocombe, K. E., and Burrows, A. M. (2014). *Primate Communication: A Multimodal Approach*. Cambridge: Cambridge University Press.

Lorenz, K. Z. (1966). Evolution of ritualization in the biological and cultural spheres. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 251, 273–284. doi: 10.1017/S0007087415000035

Marler, P. (1961). The logical analysis of animal communication. *J. Theor. Biol.* 1:295e317. doi: 10.1016/0022-5193(61)90032-7

Massen, J. J., Van Den Berg, L. M., Spruijt, B. M., and Sterck, E. H. (2010). Generous leaders and selfish underdogs: pro-sociality in despotic macaques. *PLoS One* 5:e9734. doi: 10.1371/journal.pone.0009734

Moore, R. (2016). Meaning and ostension in great ape gestural communication. *Anim. Cogn.* 19, 223–231. doi: 10.1007/s10071-015-0905-x

Moore, R. (2017). Convergent minds: ostension, inference and Grice's third clause. *Interface Focus* 7:20160107. doi: 10.1098/rsfs.2016.0107

Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biol. Philos.* 25, 755–780. doi: 10.1007/s10539-010-9224-4

Pelé, M., Dufour, V., Thierry, B., and Call, J. (2009). Token transfers among great apes (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan paniscus*, and *Pan troglodytes*):

species differences, gestural requests, and reciprocal exchange. *J. Comp. Psychol.* 123, 375–84. doi: 10.1037/a0017253

Pepperberg, I. M. (2004). "Insightful" string-pulling in Grey parrots (*Psittacus erithacus*) is affected by vocal competence. *Anim. Cogn.* 7, 263–266. doi: 10.1007/s10071-004-0218-y

Pepperberg, I. M., Gray, S. L., Mody, S., Cornero, F. M., and Carey, S. (2018). Logical reasoning by a Grey parrot? A case study of the disjunctive syllogism. *Behaviour* 156, 409–445. brill.com/beh doi: 10.1163/1568539X-00{\break}003528

Pepperberg, I. M., Koepke, A., Livingston, P., Girard, M., and Hartsfield, L. A. (2013). Reasoning by inference: further studies on exclusion in grey parrots (*Psittacus erithacus*). *J. Comp. Psychol.* 127, 272–281. doi: 10.1037/a0031641

Petit, O., Dufour, V., Herrenschmidt, M., De Marco, A., Sterck, E. H., and Call, J. (2015). Inferences about food location in three cercopithecine species: an insight into the socioecological cognition of primates. *Anim. Cogn.* 18, 821–830. doi: 10.1007/s10071-015-0848-2

Pika, S., Liebal, K., Call, J., and Tomasello, M. (2005). Gestural communication of apes. *Gesture* 5, 41–56. doi: 10.4324/9781003064541-3

Plooij, F. X., and Lock, A. (1978). "Some Basic Traits of Language in Wild Chimpanzees?," in *Action, gesture, and symbol: The emergence of language*, ed. A. Lock (London: Academic Press).

Premack, D., and Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition* 50, 347–362. doi: 10.1016/0010-0277(94)90035-3

Rossi, A. P., and Ades, C. (2008). A dog at the keyboard: using arbitrary signs to communicate requests. *Anim. Cogn.* 11, 329–338. doi: 10.1007/s10071-007-0122-3

Savage-Rumbaugh, E. S., McDonald, K., Sevcik, R. A., Hopkins, W. D., and Rubert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *J. Exp. Psychol.* 115, 211–235. doi: 10.1037//0096-3445.115.3.211

Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., and Slocombe, K. E. (2013). Chimpanzee food calls are directed at specific individuals. *Anim. Behav.* 86, 955–965. doi: 10.1371/journal.pone.0144197

Schloegl, C., Schmidt, J., Boeckle, M., Weiß, B. M., and Kotrschal, K. (2012). Grey parrots use inferential reasoning based on acoustic cues alone. *Proc. R. Soc. B Biol. Sci.* 279, 4135–4142. doi: 10.1098/rspb.2012.1292

Scott-Phillips, T. (2014). *Speaking Our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special*. London: Macmillan International Higher Education.

Scott-Phillips, T. C. (2015). Nonhuman primate communication, pragmatics, and the origins of language. *Curr. Anthropol.* 56, 56–80. doi: 10.1086/679674

Seyfarth, R. M., and Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain Lang.* 115, 92–100. doi: 10.1016/j.bandl.2009.10.003

Seyfarth, R. M., and Cheney, D. L. (2017). Precursors to language: social cognition and pragmatic inference in primates. *Psychon. Bull. Rev.* 24, 79–84. doi: 10.3758/s13423-016-1059-9

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Shorland, G. (2018). *The self and others: A glimpse into the bonobo mind*. Ph.D thesis. Switzerland: Université de Neuchâtel.

Slocombe, K. E., and Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Curr. Biol.* 15, 1779–1784. doi: 10.1016/j.cub.2005.08.068

Smith, W. J. (1977). *The Behavior of Communicating*. Cambridge: Harvard University Press.

Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge: Harvard University Press.

Tauzin, T., Bohn, M., Gergely, G., and Call, J. (2020). Context-sensitive adjustment of pointing in great apes. *Sci. Rep.* 10:1048. doi: 10.1038/s41598-019-56183-7

Tennie, C., Call, J., and Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2405–2415. doi: 10.1098/rstb.2009.0052

Tennie, C., Call, J., and Tomasello, M. (2012). Untrained chimpanzees (Pan troglodytes schweinfurthii) fail to imitate novel actions. *PLoS One* 7:e41548. doi: 10.1371/journal.pone.0041548

Tinbergen, N. (1952). Derived activities; their causation, biological significance, origin, and emancipation during evolution. *Q. Rev. Biol.* 27, 1–32. doi: 10.1086/398642

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. Berkeley: University of California Press.

Tomasello, M. (2008). Why don't apes point?. *Trends Linguist. Stud. Monogr.* 197:375.

Tomasello, M. (2009). "Primate communication," in *Key Notions for Pragmatics*, eds J. Verschueren, J.-O. Östman (Amsterdam: John benjamins publishing company), 208–216.

Tomasello, M., and Call, J. (2019). Thirty years of great ape gestures. *Anim. Cogn.* 22, 461–469. doi: 10.1007/s10071-018-1167-1

Tomasello, M., Call, J., Warren, J., Frost, G. T., Carpenter, M., and Nagell, K. (1997). The ontogeny of chimpanzee gestural signals: a comparison across groups and generations. *Evol. Commun.* 1, 223–259. doi: 10.1075/eoc.1.2.04tom

Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., and Evans, A. (1985). The development of gestural communication in young chimpanzees. *J. Hum. Evol.* 14, 175–186. doi: 10.1016/s0047-2484(85)80005-1

Tomasello, M., Gust, D., and Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates* 30, 35–50. doi: 10.1111/j.1540-5834.2005.00324.x

Tomasello, M., and Zuberbühler, K. (2002). "Primate vocal and gestural communication," in *The Cognitive Animal: Empirical and theoretical perspectives on animal cognition*, eds M. Bekoff, C. Allen, and G. M. Burghardt (Cambridge: MIT Press), 293–299.

Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., et al. (2017). Exorcising G rice's ghost: an empirical approach to studying intentional communication in animals. *Biol. Rev.* 92, 1427–1433. doi: 10.1111/brv.12289

Vail, A. L., Manica, A., and Bshary, R. (2013). Referential gestures in fish collaborative hunting. *Nat. Commun.* 4:1765.

Völter, C. J., and Call, J. (2017). "Causal and inferential reasoning in animals," in *APA handbook of comparative psychology: Perception, learning, and cognition*, eds J. Call, G. M. Burghardt I, M. Pepperberg, C. T. Snowdon, and T. Zentall (Washington: American Psychological Association), 643–671. doi: 10.1037/0000012-029

Wilson, D. (1998). "Linguistic structure and inferential communication," in *Proceedings of the 16th International Congress of Linguists 20-25* (Oxford: Elsevier Sciences).

Wilson, D., and Sperber, D. (2002). "Relevance theory," in *The Handbook of Pragmatics*, eds L. Horn and G. Ward (Oxford: Blackwell), 607–632.

Woodruff, G., and Premack, D. (1979). Intentional communication in the chimpanzee: the development of deception. *Cognition* 7, 333–362.

Yamamoto, S., Humle, T., and Tanaka, M. (2012). Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3588–3592. doi: 10.1073/pnas.1108517109

Zuberbühler, K. (2005). The phylogenetic roots of language: evidence from primate communication and cognition. *Curr. Dir. Psychol. Sci.* 14, 126–130.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

TYPE Hypothesis and Theory
PUBLISHED 02 April 2024
DOI 10.3389/fpsyg.2024.1135288

OPEN ACCESS

EDITED BY
Mark Collard,
Simon Fraser University, Canada

REVIEWED BY
Drew Rendall,
University of New Brunswick Fredericton, Canada
Lesley J. Rogers,
University of New England, Australia

*CORRESPONDENCE
Ulrike Griebel
✉ ulrikegriebel@gmail.com

RECEIVED 31 December 2022
ACCEPTED 13 March 2024
PUBLISHED 02 April 2024

CITATION
Griebel U and Oller DK (2024) From emotional signals to symbols.
Front. Psychol. 15:1135288.
doi: 10.3389/fpsyg.2024.1135288

From emotional signals to symbols

Ulrike Griebel[1,2,3]* and D. Kimbrough Oller[1,2,3]

[1]School of Communication Sciences and Disorders, The University of Memphis, Memphis, TN, United States, [2]The Institute for Intelligent Systems, University of Memphis, Memphis, TN, United States, [3]The Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria

The quest for the origins of language is a diverse enterprise, where research from a variety of disciplines brings area-specific ideas and area-specific terminology to bear. This variety often results in misunderstandings and misconceptions about communication in various species. In the present paper, we argue for focus on emotional systems as the primary motivators for social signals in animals in general. This focus can help resolve discrepancies of interpretation among different areas of inquiry and can illuminate distinctions among different social signals as well as their phylogenetic origins in animals and especially in humans. We advocate, following Jaak Panksepp, a view wherein the Seeking System, the endogenous tendency to search and explore, is the most fundamental emotional motivation. The Seeking System forms the basis for flexible, voluntary, and exploratory control of motor systems and makes much of learning possible. The relative lack of vocal learning and expression in nonhuman primates contrasted with extensive vocal learning and expression in humans began, we propose, with the evolution in ancient hominins of a necessary foundation for the many subsequent capabilities required for language. That foundation was, according to the reasoning, naturally selected in the form of neurological connections between the Seeking System and mechanisms of glottal/phonatory control. The new connections allowed ancient hominins to develop flexible, endogenous vocal fitness signals produced at very high rates and including large numbers of discrete syllables, recombinable to form syllable combinations with many prosodic variations. The increasing sociality of hominins supported evolution of massive expansion in the utilization of these flexible vocal forms to allow development of words and sentences.

KEYWORDS

origin of language, evolution of language, development of language, infant vocalization, emotional expression

Introduction

Emotions as functional motivational states: the seeking system as a foundation

Emotions are assumed by many to apply only to humans or only to animals that have highly developed brains. We advocate a view introduced by Jaak Panksepp, where at least some emotions as well as other affective states, such as hunger or the sensation of cold, are thought to apply at least to most vertebrates (Panksepp, 1998). Others have sought recently to characterize biochemical and genetic *foundations* of emotions in invertebrates such as the fruit fly (Anderson and Adolphs, 2014; Gibson et al., 2015; Anderson, 2022). Even in unicellular animals such foundations have been proposed to include conserved biochemical processes that generate exploratory foraging activity (Hills, 2006). Panksepp hypothesized that such primitive, apparently endogenous actions formed a basis for the evolution of more elaborate emotions in mammals and birds, implemented in specific neurotransmitters and largely subcortical neural circuitry.

In the present paper, we focus mostly on communication in humans and nonhuman primates (NHP), applying Panksepp's theory of affect as a basis for understanding primate vocalizations as expressions of emotion. But in other species, especially birds, calls and songs also provide evidence of emotional expression (Marler and Evans, 1996; Papini et al., 2018). For those who may object to the use of the term emotion in this way, it is possible to substitute "motivation" or "affect" for emotion. The focus is on internal states that underly certain behaviors. Our paper proposes that Panksepp's theory may provide a critical missing piece in the attempt to explain how ancient hominins broke away from the primate communication background through evolution of a powerful motivation to vocalize exploratorily, thus forming a foundation for the evolution of vocal symbology and thus, for language.

In this view, emotions and other affective states orchestrate all kinds of behavior. Emotions stimulate animals to do what is advantageous, reflecting action strategies honed over eons (Damasio, 1999). For example, when encountering danger, emotions such as rage or fear prepare mammals to take measures of attack, retreat, or freezing. Emotions also motivate mammals to reveal their emotional states vocally, providing information about what they may be likely to do, a fact that has been recognized since Darwin's characterization of emotional expression (Darwin, 1872).

Prior publications have also pointed to emotional foundations for communication and for language (Turner, 1996; Jablonka et al., 2012). These works have even emphasized "special" emotions of humans that appear to have emerged as a consequence of the extreme sociality of human groups. Panksepp also has acknowledged special human emotions that emerge from interactions of the basic emotions with cognitive capabilities and cultural adaptations that seem unique to humans.

Our paper focuses primarily on vocalization, because it is vocal language we wish to account for in evolution and development. This focus is not intended to discount the multimodality of communication in humans and other animals. Communication in primates includes facial expressions, body postures, and gestures as well as many intensity cues that modulate interpretation, which is also heavily modulated by the perceived context of communicative actions (Iverson, 2010; Taglialatela et al., 2011; Fröhlich et al., 2019). Emotions and other affective states, in accord with our proposal, are at the root of communication in all its modalities. It is also important to recognize that a dichotomous view of gesture as intentional and vocalization as emotional in NHP is not justifiable (Heesen et al., 2022b). The flexibility of usage of both gesture and vocalization in the great apes makes clear that their emotional nature does not negate their intentional and flexible usage, topics that have been richly explored in recent research (Scheider et al., 2016; Liebal and Oña, 2018; Oña et al., 2019; Heesen et al., 2022a). Research has also reported that the great apes can vocalize intentionally based on what is known by the receiver and that there is considerable contextual flexibility in how and when vocal signals are given (Crockford and Boesch, 2003; Schel et al., 2013; Clay et al., 2015; Crockford et al., 2018).

Panksepp (1998, 2005) viewed affective states, including emotions, as intrinsic motivations rather than merely as action patterns in response to environmental events. Emotions motivate flexible, targeted action in diverse external conditions, often triggered by subtle internal or external cues and often lasting over considerable periods, during which emotional states may remain activated even though the arousing event has passed. Consider how fear can last for many minutes after being triggered by an odd sound occurring in an unfamiliar place. The persistence of affective states long after the stimulus that instigates them is demonstrable in humans (Heller et al., 2015; Ivanov et al., 2020) and even in the fruit fly (Anderson, 2022).

This paper is an exploration of how emotions and other affective states such as hunger or pain motivate vocal communication, and further, how and why human vocal communication has properties that are in some ways held in common with other primates but in other ways contrast dramatically with communication in any other species. We share much of this viewpoint with other investigators (Owings and Morton, 1998; Owren et al., 2011). In some cases the connections between particular emotional states and corresponding vocal acts seem straightforward, as in aggressive sounds that can occur before or during physical conflicts. But natural selection has produced ways that animal vocal signals can be adjusted flexibly for context. Recent research has overturned previous assumptions that primate vocalization was limited to inflexible signals of emotions. For example, chimpanzees produce alarm calls in ways that are tuned to the awareness of others (Crockford et al., 2012) and are more likely to announce the presence of food if a friend is arriving than if some other individual is arriving (Schel et al., 2013). Other animals, for example, chickens have also been shown to modify alarm calls based on whether other chickens are nearby (Evans et al., 1993; Evans and Evans, 1999).

We propose that human vocalization has been naturally selected to be connected with an emotion that is typically not even treated *as* an emotion, the "Seeking System." Panksepp and his colleagues proposed that the Seeking System is the "granddaddy" of the emotions (Panksepp and Biven, 2012), motivating organisms, even in the absence of immediate conditions requiring action, to search and explore their environments. Seeking sometimes resembles foraging in unicellular animals (Hills, 2006), but things are more diverse in multicellular animals such as mammals, where search and exploration can produce not only a wide variety of experiences potentially beneficial to survival and reproduction but also a sense of pleasure to keep animals seeking even when they do not seem to need to, a point that is illustrated as fundamental in the research of Panksepp and his colleagues (Panksepp et al., 1984; Ikemoto and Panksepp, 1999; Panksepp, 2005, 2011; Burgdorf and Panksepp, 2006).

Panksepp and others have shown that laboratory rats, for example, repeatedly explore every nook and cranny of an empty box (Olds and Milner, 1954; Panksepp, 1998). They seem to enjoy this exploration, and the activity generates neuro-chemical pleasure signatures (especially dopamine). Rats with an implanted electrode in a Seeking System tract, the medial forebrain bundle of the lateral hypothalamic area, have been found to repeatedly self-stimulate by a lever press and explore until they are exhausted. Seeking System activity is described as including "…intense and enthusiastic exploration and appetitive anticipatory excitement …" (Panksepp, 2010, p. 537).

Mammals show a deep motivation to understand the world, to watch, touch, listen, taste and smell, being motivated by curiosity and interest, a condition that regularly occurs while they are awake, looking for novelty as illustrated in the research of Panksepp and colleagues (e.g., Panksepp and Biven, 2012). Seeking can also be thought of as driving planning, strategizing, and imagining. Humans appear to seek information through travel and reading, and the search for entertainment can also be viewed as motivated by the Seeking System, although we know of no experimental work with humans to illustrate this apparent tendency. The neurological and neuro-chemical foundations of the Seeking System were described in a variety of animals through extensive research reviewed by Panksepp (1998, 2005, 2011) and Panksepp and Biven (2012).

The Seeking System is not present in traditional emotion categorization schemes, perhaps because behavioral research has often been limited to consideration of "stimulus and response." The endogenous, active organism in Panksepp's (1998) theory constitutes a dramatic departure from strict behaviorism, because so much of what animals do is *not* seen in his emotion-centric conception as reactive to the external environment, but as the internally-generated behavior of an organism motivated to explore. Panksepp (1998) and Panksepp and Biven (2012) acknowledged repeatedly that behavior is responsive to the external environment, but they argued that animals seek information regardless of external circumstances. They referred to the work of Olds and Milner (1954), who had concluded that electrical stimulation of the septal area "is rewarding in the sense that the experimental animal will stimulate itself in these places frequently and regularly for long periods of time if permitted to do so (p. 419)."

We contend that vocal communication in most primates is predominantly motivated by emotions that are tied to *immediate needs* and triggering events. This interpretation is consistent with the original Darwinian description of emotional expression (Darwin, 1872). It does not appear that NHP *vocalizations* generally require involvement of the Seeking System, although the Seeking System does, in accord with Panksepp's (1998) theory, motivate NHP to develop intelligent, forward-looking strategies for foraging, hunting cooperatively, warfare with neighboring troops, alliance formation to overthrow tyrannical alphas, and so on (de Waal, 2016, 2019). Vocal communication instead appears to be motivated in NHP by emotions of the here and now, such as fear or rage and others to be considered below in fuller discussion of Panksepp's theory.

Human vocalizations are sometimes also motivated by immediate emotional triggering events, for example, by fear or rage, but the great majority of adult human vocalizations do not require such triggering. Human vocalizations regularly merely require, in accord with Panksepp's model, motivation through the Seeking System. We propose the Seeking System motivates human infants in the first months of life to produce ~3,500 speech-like vocalizations or "protophones" per day (Oller et al., 2019, 2021), the vast majority of which are directed to nobody (Long et al., 2020), but seem to constitute endogenous exploration of the vocal capacity (Oller and Griebel, 2021). We propose further that the connection of vocal control to the Seeking System in hominins established first in ancient hominin infants.

In accord with Panksepp's theory, we propose that because of the connection of the Seeking System to human vocalization, we can explore vocalization and at other times vocally express any emotional state flexibly with those same sounds, and in maturity, with an indefinitely large set of words and sentences. The Seeking System, in our interpretation, allows human vocalization to require no targeting of immediate benefits, and thus frees the human vocal capacity for expressions that can bear any social function or "illocutionary force" (Austin, 1962; Oller, 2000), a concept elaborated below. Ultimately human vocal expressions become symbols that can refer not only to entities in the here and now, but to entities in the past, the future, or the imagination. Our proposal is that the connection of human vocalization to the Seeking System makes these advances beyond NHP communication possible.

How is language possible, and indeed how is nonhuman animal communication possible?

Biologists have often wondered how communication signals could have evolved at all, since so much of communication appears to be about conflict, competition, and deception (Zahavi, 1975; Krebs and Dawkins, 1978; Knight, 2016). Wouldn't signals deteriorate before they could stabilize because of conflicting interests? Maynard Smith and Harper (2003) argued that signals stabilize because in general they benefit both the sender and the receiver, conveying useful information about the sender's affective state. If a baby communicates distress because of hunger or fear, it is in the interest of the mother to react by giving care, since the promotion of her genes is at stake. The benefits of this type of communication can be extended to relatives in general (Hamilton, 1963) as well as to unrelated group members who might return the favor later in a tit-for-tat arrangement (Trivers, 1971).

One might expect aggressive signaling to give the sender the edge over the receiver, but aggressive signaling is also important for the receiver, who can benefit from knowing how motivated the sender is. The receiver might wisely choose to back off after perceiving a highly motivated aggressor. Protection of home often wins over threatened attack because the home protector may have more to lose and have higher motivation to fight (Hoefler, 2002). In general it is to the advantage of both parties in aggressive exchanges to display their motivation regarding a possible fight and to assess the motivation of the opponent.

Much research has addressed cheating and manipulation in such communicative acts as reviewed by Oller and Griebel (2021, this Frontiers Topic). We propose that communication systems are evolutionarily stable in part because animals display and assess each other's affective states and benefit from the interchanges in guiding their actions. The existence of deception does not contradict this view. In stable communication systems, deception has to be rare, because otherwise signals would disintegrate—they have to be generally reliable indicators of emotional states and the related intentions. Experience in social groups can also play a major role in maintaining stability of shared signals. Group members tend to learn whose emotional signals and communicative intentions to take seriously. For example, inexperienced youngsters' alarm calls may be ignored (Cheney and Seyfarth, 1988). Thus, vocal signals must be recognized by animal receivers as revealing affective states, but their responses can be intelligent and flexible based on what they know about the sender, the sender's perceived intentions, or the situation.

NHP vocal signaling appears generally to elicit responses presumably consistent with innately determined purposes of the sounds, such as maternal caregiving in response to infant distress calls. The same is true in the human case during early infancy, but across development the producer's intended function and the listener's interpretation of crying can change because human infants learn to cry manipulatively rather than only as an expression of distress (Green et al., 2011; Chóliz et al., 2012). By human adulthood, crying and laughter can even be expressed in ways that are contrary to their innate purposes. Think of derisive laughter or crying in relief. We propose that this flexibility and the seemingly unlimited emotional flexibility of linguistic expression are possible in humans because all vocalizations in humans can be motivated through the Seeking System, which allows vocal flexibility to an extraordinary degree.

Understanding the relation between the Seeking System and human vocalization may provide a key to clarifying both similarities and differences between human and NHP vocalization. If we focus on the Seeking System, we may be able to sort out many prior difficulties of interpretation that have led animal communication research down unproductive paths, directing attention away from fundamental differences between vocal communication in humans and nonhuman animals, while also failing to address the most salient similarities. We propose that the relations between vocal communication in humans and nonhuman animals need to be restructured in both conceptual foundations and terminology, taking account of a role for the Seeking System. The suggestion that the Seeking System may have come to be connected to a far greater extent with vocalization in humans than in NHP is consistent with the fact that humans have extensive neural connections between laryngeal motor cortex (LMC, located in area 4 of the primary motor cortex) and laryngeal motoneurons, while NHP have been shown to possess little if any such connection (Jürgens, 1995; Kumar et al., 2016; Simonyan et al., 2016). Laryngeal motor neurons in monkeys appear to have indirect connections from area 6 of the premotor area, but their destruction does not appear to impact species-specific calls (Simonyan et al., 2016). Consequently it seems possible that the Seeking System in humans (perhaps from its lateral hypothalamic site) may be shown to have strong connections with LMC.

Explaining discrepancies in the comparative literature on vocal communication

Interpretive and terminological misunderstandings about signaling repertoires

One vexing issue in description of signaling systems is that the numbers of communication signals in species repertoires have been reported to be vastly different. For example, in a broad review of primate communication literature, chimpanzees were reported to possess from 7 to 27 different vocal types across studies, and similar discrepancies were found for other primate species (Sutton, 1979). Obviously, the criteria for counting differed across the studies reviewed.

Perhaps the main reason for such discrepancies is that signals along a particular dimension of emotional expression grade from low to high intensity, and they can show regime shifts of sound quality along each such dimension. Squirrel monkey vocalizations, for example, have been reported to pertain to five functional dimensions (protest, challenge, social contact, group action, and alarm), with gradations on each one corresponding to apparent regime shifts, yielding the misleading impression of many more than five distinct functional categories (Jürgens and Ploog, 1976; Ploog, 1992). Interestingly, at very high intensity, all the emotional dimensions of squirrel monkeys tended to collapse acoustically, yielding a single loud, dysphonated sound. At very low intensity, all the dimensions seemed to collapse to a single very quiet sound, so that middle range intensities yielded best discriminability (Jürgens, personal communication; and see Oller, 2000, pp. 339–355).

Especially in very social animals such as primates, it is important for receivers to assess the intensity of the emotions driving signals. This intensity is reflected in graded vocalizations, with seemingly categorical shifts, especially with sudden changes from periodic phonation at low intensity to noisy dysphonation at higher intensity (Owren and Linker, 1995; Owren and Rendall, 2001; Rendall and Owren, 2013). The sender's degree of motivation can also be reflected in the number of repetitions of vocalizations.

Failure to take emotional motivation into account appears to have contributed to the narrative in much animal communication literature portraying animal vocalizations with linguistic terms. The foundational research on this topic was based on observations of vervet monkeys, *Chlorocebus pygerythrus* (Seyfarth et al., 1980). The research inspired the widespread claim that various animals have "semantic" vocalizations in the form of predator-specific alarm calls (Seyfarth et al., 1980). The term was misleading, as will be seen below, having been largely rejected by more recent investigators (Owren et al., 2010; Rendall, 2021) and having been substantially weakened in more recent time even by the originators of the claim (Cheney and Seyfarth, 1996). Even so, additional researchers have interpreted distributional differences in the acoustic characteristics of alarm calls of a variety of species in the presence of different predators as indicating "referentiality," "functional referentiality" or "representation" (Gozoules et al., 1995; Hollén and Manser, 2007; Furrer and Manser, 2009; Rogers et al., 2018, and see a variety of papers in a volume by Stegmann, 2013).

Yet the term "predator-specific alarm calls" of vervets, which lay at the basis of the semanticity claim, was misleading all along, since it has been widely acknowledged that the relevant calls are not specific to predators. They are also commonly used in intra-specific aggression, that is, the same calls occur during fighting and threats by one vervet in conflict with another (Price et al., 2015). This fact was even noted briefly in the article that first presented the claim about semantics in vervet monkey calls (Seyfarth et al., 1980). The confusions about alarm calls are placed in historical perspective in two additional articles in this Frontiers Topic (Locke, 2021; Rendall, 2021).

The tendency for calls associated with predator danger to segregate into acoustically different groupings can easily be interpreted in a way that requires no appeal to semantics. Instead, the segregation can be interpreted as relying on emotional expressions occurring differentially because of the different intensities of fear or anger that can occur at typical sighting of particular predators (Oller and Griebel, 2014). Different degrees of fear and anger can mix in ways that are appropriate for particular situations and may yield different sounds in response to a predator that may typically be seen crouching on the ground as opposed to one that may be soaring in the air or slithering in the grass. The extent of the alarm-caller's emotional reaction may segregate probabilistically so that one type of predator tends to cause more alarm than another simply because they may tend to be nearer when detected or may tend to be approaching faster when detected.

Given the existence of regime shifts in graded vocalizations, it seems possible that typical alarm calls to one type of predator might misleadingly seem categorically different from those to another, only owing to probabilistic differences in the typical intensity of the reaction.

It has not even been proven that the producer of a so-called alarm call *intends* the vocalization to constitute an alarm. A human observer may think of the sound as an alarm, and conspecifics of the caller may respond *with* alarm, but we know of no evidence indicating that the animal produces an alarm call based on anything other than fear and/or anger. An increased tendency to produce alarm calls when kin are nearby could be interpreted as simply indicating the sender feels more fear or anger knowing their kin might be in danger. Of course it is an empirical question what is in the mind of an animal signaler (long-term plans and thoughts could indeed be involved), but an alarm call itself does not include semantic information about the mind of the signaler. For the receiver, it only supplies emotional/illocutionary information, reflecting the caller's state and immediate intentions. Other concerns about interpretation of alarm calls were expressed by Kaplan (2008) based particularly on evidence from the Australian magpie. Her conclusion was that the various calls that have been termed "referential" appear to be generated principally in the midbrain, offering little support for any interpretation of complex cognition being involved. Still, while an animal signal itself may reflect only the type and intensity of the sender's emotion/illocution, the listener may bring to bear contextual information and prior experience in determining how to react.

Different distributions of animal food calls in the presence of different edibles have also been interpreted as referential or functionally referential (Evans, 1997; Evans and Evans, 1999; Rogers et al., 2018), but in our opinion this interpretation is subject to the same concerns as the interpretation of animal alarm calls. The empirical evidence does not appear to prove that the calls themselves contain reference. As in the case of alarm calls, the possibility remains that the differences among the calls in the presence of different kinds of food are the products of different emotional reactions of the vocalizers to the different food types. This does not mean, however, that the food calls have no flexibility since, for example, their production has been shown in some instances to reflect audience effects (Hauser and Marler, 1993; Evans and Evans, 1999; Schel et al., 2013). Nor does the possible lack of reference contained within the calls themselves rule out conspecific listeners' reactions being differentiated based on possible learning by the community of listeners about the likelihood that different foods may have stimulated the senders' differentiated vocal reactions.

Our conclusion is that neither food calls nor alarm calls in animals have thus far offered evidence of semantics. For any vocal act to be semantic it is required that it be motivated by a wide variety of states and intentions, for example to involve simple naming in the absence of alarm or in the absence of any kind of food and in the absence of any particular emotional state. Human linguistic reference is *never limited to* specific circumstances of physical context or emotional state. The failure so far to produce convincing evidence of semanticity in animal communication in the wild does not appear to be a failure of methodology or of investigator persistence. It appears instead to be a failure due to inappropriate goal-setting—researchers have sought to show advanced features of human language (especially semanticity) in wild animals without first taking stock of the fundamental principles of cognition and behavior that are required by such language features. The same researchers have largely left aside the investigation of fundamental *differences* between human and nonhuman animal communication.

There are, however, ways to compare human and nonhuman animal communication profitably. To find both similarities and differences between them, we can look to non-linguistic human communication modes that have much in common with nonhuman animal vocalization, such as the human non-verbal vocal repertoire, facial expressions, and non-symbolic gestures that are clearly associated with emotions or other affective conditions. Human infant crying (which continues intermittently in maturity in modified and much more flexible form), for example, has much in common with the calls of other mammals, since crying expresses distress caused by pain, fear, or isolation (Owings and Zeifman, 2004). Similarly, laughter has been interpreted as occurring across many primates, and although the acoustic form of laughter differs across species, it seems clear there is homology involved (Davila Ross et al., 2010). As with crying or screaming, the cross-species similarity in laughter is grounded emotionally—laughter's central function is always the emotional expression of social connectedness or joy, which in NHP tends to occur in response to tickling or rough and tumble play (Panksepp, 2000).

Laughter and crying have sometimes been referred to as "fixed signals" (Lorenz, 1951; Tinbergen, 1951), but this is an overly restrictive term, because these sounds can show substantial gradations of intensity. Laughter and crying are indeed relatively fixed in that they each have a limited range of variability regarding their functions, either the expression of a positive social emotion or the expression of distress. In this way human and nonhuman animal vocalizations have something fundamental in common.

In contrast, even the precursors to language seen in human infancy are not constrained to expression of particular functions the way crying and laughter are. On the contrary, from the first months of life, human infants express the full range of affective valences from positive to neutral to negative with each of the phonatory vocal types (e.g., squeals, vowel-like sounds, growls) known to be precursors of speech (Oller et al., 2013; Jhang and Oller, 2017). Infants can shift in just a few moments between emotional states, and they can use any one of their vocal types to express any emotional valence. Thus one may observe an infant to produce a squeal with a smiling, happy face, then later a squeal with a neutral face, suggesting neutral emotional valence, and later yet with an obvious frown indicating discomfort or annoyance. Human observers make consistent judgments about the differing emotional states accompanying the very same vocal types on different occasions (Oller et al., 2013). All elements of mature human languages require this kind of functional/illocutionary flexibility—every syllable, word or sentence must be possible to produce in any emotional state. In fact, each element of language must be possible to produce merely based on interest in the sound itself. Infant protophones produced this way are judged consistently by human observers to constitute exploration (Long et al., 2020; Oller et al., 2021), and we propose this exploration depends on connection to the Seeking System.

The endogenous nature of human infant vocal exploration is indicated partly by the facts that (1) the great majority of protophones are directed to nobody (Long et al., 2020), (2) they express a variety of emotional states (Stark, 1981; Shimada, 2012), (3) much of the infant vocalization occurs when infants are alone in a room (Delack and Fowlow, 1978; Oller et al., 2019), and (4) even profoundly deaf infants produce massive numbers of protophones (Iyer and Oller, 2008). Most protophones seem to constitute a kind of vocal play (Stark, 1980), a fact that is supported by observations from human coders of hundreds of randomly-sampled segments from all-day recordings across the first year of life as well as from extensive longitudinal research with laboratory-based audio-video samples (Oller et al., 2021; Long et al., 2022). Similarly, adult humans produce speech commonly (often muttering to themselves) for no immediate social purpose, perhaps as a sort of anticipatory practice, expressing interest in both the speech itself and the semantic content it could at some point transmit to an imagined listener. Yet even recent efforts to explore the possibility of vocal functional flexibility in nonhuman animals (Clay et al., 2015; Dezecache et al., 2020; Taylor et al., 2022, 2023) have not reported nonhuman apes producing any vocalization exploratorily. So far neither vocal exploration nor semantic communication has been found in NHP vocal communication, and it would be more fruitful to focus on similarities and differences between signal types that, like crying and laughter, are in general based on similarly constrained functions.

The inappropriate attempt to shoehorn the vocal communications of nonhuman animals into categories of human language is exacerbated by failure in much animal communication research to draw the critical distinction between situational context and function of communication signals. Some research has considered *only* situational context in categorizing signals, ignoring social functions entirely. For example, the term "food calls" or "food associated calls" (Hauser and Marler, 1993; Clay and Zuberbühler, 2009) suggests that calls produced near food or in anticipation of eating are *about* food. It is much more plausible that they may be expressions of positive excitement or of some other emotional state that can be expressed in the absence of food. Other research by some of the same authors makes clear that vocalization in the presence of food by mammals and birds does not provide referential information (Clay et al., 2012).

The same behavioristic approach can lead to aggressive or fear expressions being categorized as predator-specific alarm calls, because they *can* occur in the context of sighting a predator, although they also occur in intra-specific aggression. NHP signals can be used in a variety of different contexts, because communicative functions, such as aggression, appeasement, courtship, play, and fear, are transmitted in contexts as different as resting, traveling, and feeding. Categorizing signals by situational context alone, instead of also addressing social function, creates confusion in comparative communicative research. Social functions/illocutions of the vast majority of NHP vocalizations can be best explained as being driven by emotions that were naturally selected to serve particular social functions in the immediate present. It has not been proven that NHP vocalizations ever express any kind of semantic content, which requires by definition flexible control and learning of conventional symbols.

Roots of confusion about language and nonhuman animal communication in radical behaviorism

The confusions associated with trying to categorize animal vocal communications in terms of human language categories have also been exacerbated by a longstanding tendency in the study of animal behavior to focus on *and only on* observable, countable events, and to ignore explicitly the internal states of animals and motivations that drive behavior (Watson, 1913; Skinner, 1957). Situational context has been at center stage in radical behaviorism. Internal states of organisms have been treated as irrelevant. But there is no accounting for how and why behavior occurs without addressing the states within animals that are the immediate and necessary drivers of behavior. Situational context can help us to *infer* internal states, but to explicitly ignore internal states is to abrogate the ultimate responsibility of ethological science (Lorenz, 1971), which is to explain behavior and its evolution.

One of the results of the tendency to confine behavioral science to observables is that scientists have often seemed to take deeply contradictory stands on whether animals possess emotions at all (see commentary in Panksepp, 2005; Panksepp and Biven, 2012). Biologists who have actually wanted to study emotions during the century of the behaviorist paradigm have sometimes complained that animals are treated as simple stimulus–response machines without any emotions or minds, while psychologists trying to explain human behavior may complain that humans struggle with their emotional heritage, their lowly "animal" side, which is thought to run afoul of the humans' highly evolved rational cognitive abilities. We tend to agree with Panksepp (1998) that there can be no cognition without affect, that there can be no learning without emotional motivation, and that emotions are modulated by learning and memory in both humans and many other animals. In this view, emotion and adaptation by learning had to evolve together, like two sides of a coin.

While visual and chemical signals are also used to communicate emotional states, the vocal mode is naturally connected to the breathing apparatus in vertebrates, and consequently to arousal, which can reflect levels of emotions through respiratory pressures, volumes and rates. Phonatory mechanisms rely on respiration as the force to drive vibratory patterns of larynx, syrinx, pharyngeal pouches, air sacs, etc. (Conrad and Schönle, 1979; Kent, 1998; Suthers, 1999; Farmer, 2006; Schusterman, 2008). Vocalization can thus occur to some extent accidentally in circumstances of high arousal and intense emotion. Consequently we surmise that vocalizations, being naturally associated with affective states through breathing, have been particularly sensitive to selection pressures that could have differentiated vocalizations to express individual affective types and for elaborating them into graded signals.

Visual systems of communication can be flexible and can include relatively large repertoires, as in the case of the chromatophore system in cephalopods (Messenger, 2001; Byrne et al., 2003) or gestural communication in NHP (Call and Tomasello, 2007; Call, 2008). But with the exception of human sign languages (Stokoe, 1960), only vocal sound production has produced massively complex categorical signal types, and these have occurred only in human languages and in certain species of birds and marine mammals showing elaborate songs with an amazing variety of "syllable" types (Helweg et al., 1992; Marler and Slabbekorn, 2004).

Emotions at center stage

Panksepp's view of emotions

While widely recognized models of emotional systems (Ekman, 1994; LeDoux, 1998; Damasio, 1999; de Waal, 2019) show substantial overlap with Panksepp's (1998) model, they also differ importantly. LeDoux (1998), for example, has resisted even acknowledging that emotions exist in nonhuman animals, and in general other models do not provide a basis for characterizing the origins of vocal emotional expressions in either humans or nonhuman animals. Panksepp pioneered characterization of emotions and other affective states (including hunger, thirst, pain) both theoretically and in neurobiological research.

While all of the other basic emotions of Panksepp's model to be considered below are commonly recognized in some form, the Seeking System is an essentially new concept. There has been mention of an exploratory "drive" that causes animals to investigate their environments, but the description of this "drive" prior to Panksepp was vague and largely undefined neurologically (Montgomery, 1954; Glanzer, 1958). Recent modern computational modeling research includes much interest in curiosity and exploratory behavior (e.g., Oudeyer and Smith, 2016), but Panksepp (1998) and Panksepp and Biven (2012) have supplied the most extensive characterization of the Seeking System.

Panksepp described the Seeking System as the most fundamental emotional system, motivating all animals to explore. The Seeking System presumably also plays a crucial role in innovative behaviors such as tool use, and we have proposed specifically that it has been crucial in the evolution of symbolic communication/language, since exploration of vocalization is one of the fundamental foundations for language. Without it, other prerequisites of language could not have emerged. The argument about additional stages of evolution and development that require the foundational stage of vocal exploratory flexibility can be found in our previous writings (Oller, 2000; Oller et al., 2016), but for the present, we simply note that the following list of four capacities appear in human infants in the following order across the first year:

(1) exploratory production of vocal types that are not part of the innate repertoire (e.g., cry or laughter) along with production of those vocal types in any condition of affect; vocal types occurring in exploration, presumably motivated by the Seeking System, can be said to be decoupled from any of the traditionally recognized emotions;

(2) flexible, sustained, affectively positive vocal interaction with others, forming a foundation for understanding others' affective states and minds;

(3) vocal imitative learning of new sounds not in the innate or exploratorily developed vocal repertoire; and

(4) associative learning of the production of acquired sounds in association with arbitrary circumstances or entities, i.e., primitive word learning.

All four of these prerequisites to symbolic word learning (Sinha, 2004) regularly appear developmentally in the order given in humans, where #2 depends on #1, #3 and on #2, and #4 on #3. There are a number of additional steps specified in the more elaborate version of our model (Oller et al., 2016). Our key contention is that when the Seeking System was naturally selected to be connected to the human vocal system, and thus exploratory vocalization began in ancient hominins, a door was opened to evolution of many subsequent language-necessary capacities. Similarly, modern human development begins with exploratory vocalization from the first day of life, driven, according to our proposal, by the Seeking System.

Panksepp has described the neural circuitry and neurotransmitters active in seven basic emotions (Panksepp, 1998, 2005, 2011). All are portrayed as largely subcortical in vertebrates, although all can be modulated by cortical influences and by interactions across the different emotions. It is important to keep in mind, though, that to date science has produced minimal knowledge about a large number of only recently recognized neurotransmitters found in the neuronal synapses (Grant, 2015; Zhu et al., 2018). While major types have been identified (oxytocin, serotonin, dopamine, estrogen, testosterone, and so on), we know very little about how they interact with each other and with all the other neurotransmitters about which virtually nothing is known. The picture is getting more complex as research shows that the effects of a single neurotransmitter can be diverse and highly time-and area specific (see, e.g., Young, 2012). Nevertheless, there is reason for optimism since neurophysiologists have begun to characterize emotion-like systems in model organisms such as fruit flies and mice, starting from behavior and continuing to neurotransmitters and genes involved in the expression of primitive emotional states (Tsien et al., 1996; Adolphs and Anderson, 2018; Jung et al., 2020). Panksepp postulated that together with sensory/perceptual input, seven basic emotions create a primary process consciousness in mammals, which can be elaborated into states of secondary and tertiary consciousness, at least in humans, by reflections about experiences and reflections about reflections (Panksepp and Biven, 2012).

Summary of the seven emotions proposed by Panksepp

The first emotion in Panksepp's (1998) theory may apply even to unicellular animals, but at least to multicellular animals:

1. Seeking: Panksepp described this addition to the traditionally recognized emotions as an affective state of exploration, a dopamine-driven seeking/expectancy/wanting system that energizes activities such as foraging or object exploration/play (not social play, which belongs to the Play emotion, below) and mediates anticipatory states. Seeking stimulates a positive reward system (demonstrated especially in rats, where his research was most extensive) different from rewards associated with other emotions. There are no specific vocalizations associated with seeking, but if the Seeking System is connected to the vocal capacity, as we propose it is in humans and in other vocal learning species such as some song birds and some marine mammals, then an indefinitely large range of sound types and gradations of each type can emerge as a result of seeking/exploration.

Three emotions are common to both social and non-social vertebrates and perhaps also invertebrates:

2. Rage: a state of anger often expressed vocally by growling, roaring, barking, or hissing, depending on the species.
3. Fear: a state of negative agitation, yielding responses of freezing or flight, often accompanied by whimpering or screaming.
4. Lust: a state of sexual interest, yielding mating behaviors, with vocal expressions sometimes including sighing, moaning, or other sounds suggesting positive arousal.

The third group of emotions is unique to social animals:

5. Panic/Grief: a state of sadness or terror based on isolation often leading to frantic search or, after extended periods, depression.

A vocal expression is isolation calling, but after a longer period of isolation, vocalizations can include sobbing or wailing.

6. Care: a state inducing enjoyment in investing in the well-being of others, usually offspring, but also sexual partners, and allies. Behaviors include nurturing, cuddling, helping, teaching, and so on. Vocal expressions can be soothing sounds at relatively low intensity or celebration sounds of positive arousal.
7. Play: a state of pleasurable social connection in seemingly (i.e., momentarily) unproductive behaviors such as, tickling, chasing, and rough and tumble play, often accompanied by laughter.

These seven are the emotional affects. Bodily regulatory urges/states such as hunger, thirst, the need to defecate, and so on, are categorized as "homeostatic" affects in Panksepp (1998, 2010) work, and the pleasures and pains of externally provoked sensations such as sweetness, bitterness, heat, coldness, or physical injury, are treated as "sensory" affects. These non-emotional affective states can also yield vocal expressions, such as infant crying with hunger or pain or sighing in response to a pleasurable taste.

Panksepp (1998) described various inhibitory and excitatory interactions among the basic emotions. Behavioral evidence from mammals (including humans) in addition to introspection show that we can experience mixed emotions (de Waal, 2011; Larsen and McGraw, 2014; Hoemann et al., 2017). We can be torn between fight or flight, love or hate, and we often seek rational solutions to emotional conflicts, a state that Panksepp reasoned to often invoke the Seeking System in order to acquire information needed to create a balancing strategy for action. Seeking can recruit memories of prior experience acquired during prior seeking and can support informed strategic action in response to conflicting emotions. Panksepp also contended that the Seeking System can be recruited to serve the goals of other basic emotions such as Fear (seeking an escape route, a means of defense, etc.), Lust (seeking ways to impress a potential sexual partner), Rage/aggression (seeking ways to impress an opponent, to get the upper hand in a physical fight), and so on. Thus emotional states are not completely isolated from each other, although they all, according to his research, have isolable subcortical components and biochemical signatures in mammals. Interactions are obviously necessary since, for example, Fear has to interact with other emotional systems to produce reaction strategies. Everyone knows the feeling of being torn between fight or flight, and we recognize the fierce and seemingly fearless aggression of a mother with pups (in this case Fear is suppressed, presumably by Rage), the balance of Fear and Lust in courting, and the tabling of behaviors based on prior emotions in the face of imminent danger. In the case of danger, Fear may dominate every other emotion. Since vocalizations are motivated by affect/emotion, the occurrence of mixed emotions suggests that the apparent range and number of possible vocalizations may be increased substantially beyond that which would be expected if each vocalization type corresponded one-to-one to a single emotion/illocutionary type.

Reward systems in communicative behavior

Just as adaptive behaviors like reproductive acts, parental care, social bonding, play, and exploration often produce pleasure, communicative acts also have a pleasurable component. They appear to release endorphins and boost our immune functions, e.g., via opioids (Benson, 2019). While pleasure is obvious in most cases of laughter during social play (Manninen et al., 2017) in humans and our primate relatives (Davila Ross et al., 2010) and in some other more distantly related species (Panksepp, 2000), it seems likely that courtship displays such as mating songs or territorial or social group choruses in, for example, canids and primates induce pleasure as well. It has been shown that singing in a choir releases endorphins in humans (Launay and Pearce, 2020).

Panksepp's (1998) extensive review of evidence that seeking is pleasurable suggests exploratory vocalization could yield pleasurable sensations, just as exploration by smell or by touch in a variety of animals could yield pleasure. Indeed, we propose that one of the proximal mechanisms sustaining exploratory vocalization, occurring at massive rates in human infants and presumably continuing as solitary muttering in adults, is that, at least in part, it is fun, with endorphins being main ingredients of the sensations.

From emotional signal to emotional language

Differentiating features of language from features of nonhuman animal vocalization

A key difference between NHP vocal communication and an act of language is that a NHP vocal signal is coupled to a specific function or class of functions for which it was evolved. We know of no report of a NHP using a specific aggressive signal (e.g., a growl) in any non-aggressive state, except in play, where many behaviors can be used to "pretend." Similarly, courtship signals require a state of Lust, and a call for help appears to require a state of Fear.

Language, on the other hand, in its mature form, consists of conventional and learned symbols, and these symbols can signify entities (objects, actions, states, and so on) abstractly, with no necessary connection to a social function. A language act or "Illocution" (Austin, 1962) is motivated by emotions, of course, but we can use any symbol or symbol sequence of language to express any emotional state and thus also a seemingly unlimited variety of illocutions. For example, we can use the word "mouse" in various emotional states, and thus we can transmit various illocutionary forces with the word. Someone might, based on Fear, jump onto a chair and exclaim in an illocution of alarm, "a mouse!." A child opening a birthday present, being delighted to find a live mouse, might express an illocution of celebration, saying excitedly "a mouse!" A teacher pointing to a chart displaying rodents might produce a labeling illocution, saying "a mouse." "Little mouse" (Mäuschen) is used in Austria as a popular illocution of endearment spoken to both children and lovers. We have used "mouse" in this paragraph as an example, with the motivation to exemplify/share information about the illocutionary flexibility of words in languages. Importantly, the word "mouse" does not change its meaning (the class of animals it makes reference to) in these different illocutionary implementations and in the different emotional states that might have motivated them; the word always refers to a special kind of rodent that we are all familiar with and that we agree implicitly to call "mouse" in the English language.

This distinction between illocutionary function and meaning (or semantics) is critical to the understanding of a fundamental difference between language and the natural communication systems of NHP and other nonhuman animals. Wild NHP transmit illocutionary forces when they vocalize or gesture, but there has been no demonstration that they transmit semantic content. Humans can communicate in some cases with illocution alone, especially with their non-verbal repertoire (grunts, screams, cries, squeals, moans, and so on), and illocutions provide the only form of communication in the human infant. But at later stages of development, linguistic utterances appear, and these utterances include semantic content, in the form of words, generally referring to entities that need not be present. Yet every use of a word can bear any one of a large number of illocutionary forces. As in the example above, "mouse" always refers to a class of animals, but in any individual speech act, the word can constitute an alarm, an insult, a threat, an endearment, an act of labeling, or merely a pronunciation of the word.

In the first year of protophone usage in infancy, humans produce a narrow range of possible illocutions, limited to expression of just a few emotional/affective states, for example anger, distress, delight, fear, comfort, and sound exploration (vocal play) (Papaeliou et al., 2002; Scheiner et al., 2006; Oller et al., 2016). Similarly, we see no reason to believe that the range of possible illocutionary forces in NHP having grown up without human training in communication is much different from the human infant range. The list of possible illocutions appears to be confined more or less to expressions of anger (threat), fear, distress, delight (celebration), contact, submission, and perhaps a very few more (Griebel and Oller, 2008). The list should be expanded to include expressions of mixed emotions/illocutions. Saliently absent in the list, however, is vocal exploration, which to our knowledge has never been reported to occur in any of our ape relatives.

The human list of possible illocutions expands vastly as language expands through syntactic constructions that combine words into sentences. Thus we become able to express illocutions that are impossible even to imagine in nonhuman animal communication systems. These illocutions include labeling, requesting, thanking, welcoming, description, criticism, praise, denial, affirmation, argument, explanation, stipulation, and many more. Illocutionary types tend to be restricted to expressing particular emotions in nonhuman animal communication (for example, threat goes with Rage, alarm goes with Fear), but in humans there are illocutions that can be motivated by any emotion. Consider an explanation. A person can give an explanation: (a) out of mere interest in exploring an idea with someone (Seeking); (b) to counter an insulting accusation made by someone (Rage); (c) to prevent someone from striking out based on a misunderstanding (Fear); (d) to provide a rational basis that someone might like to engage in courtship behavior (Lust); (e) to provide a basis for forming a friendship (Care); (f) to justify a playful wrestling activity (Play); or (g) to evaluate irrational fears or feelings of isolation (Panic/Grief). In this way the connection of vocalization to the Seeking System in humans makes it possible not only to form a vast number of signal types, but also to use those signal types with a seemingly unrestricted variety of social intentions (illocutions), motivated by any emotional or affective state.

It may be important to emphasize that NHP appear also to transmit some illocutions on the basis of different emotions on different occasions. In gesture, for example, an invitation can be made to play or to groom (Fröhlich et al., 2016; Heesen et al., 2021). In terms of Panksepp's (1998) scheme, it would seem that a play invitation would be motivated by the Play emotion, and a grooming invitation by the Care emotion.

Humans can, by recombination of illocutionary types expressed in complex syntax, produce a large number of illocutions in a single sentence. For example, a mature language user can form a sentence constituting a request for an affirmation of an explanation of a request. Because we can, in this way, embed various illocutions in complex sentences, there is no obvious limit to how many complex illocutions are possible in human language.

The (theoretically) indefinite size of the repertoire of human illocutions reflects the similar indefinite size of the class of possible sentences that can be composed over any human vocabulary by recombination and structural embedding of phrases consisting of words (Chomsky, 1965). In fact, without complex syntax, expression of complex illocutions is not possible. And the words of which sentences are composed are themselves sequences of syllables that can be recombined to form an indefinitely large vocabulary of words (Pinker, 1994).

The recombination of discrete units at all these levels (illocution, syntax, vocabulary) is based on a digital rather than analog system that has been related to the "particulate principle" in the organization of both inorganic and organic systems. Importantly, the particulate principle in human language (Studdert-Kennedy, 2000, 2005), operating according to our reasoning on the basis of the Seeking System, makes the repertoires of language at every level vastly different from the comparatively tiny set of vocal communication types available to NHP. A key difference is that NHP vocal types seem to be obligatorily analog, offering gradations along a few dimensions of illocution but with no particulate digitization necessary for recombination that would allow indefinitely large sets of communication units.

Figure 1 portrays perhaps the most fundamental difference between human and NHP vocal communication and corresponds to the first item on the list above of 4 ordered steps in human development. NHP signals tend to present couplings between individual action types (signals themselves) and individual social functions (or illocutions) motivated by particular emotions (or complex illocutions motivated by mixed emotions). Recent research suggests there is some, but quite limited, flexibility in these couplings in bonobos and chimpanzees (Clay et al., 2015; Taylor et al., 2023). In the human case, on the other hand, the flexibility is extreme. Every language signal type and even each type of protophone precursor to human language has no necessary coupling with any particular illocution or with any particular emotion. Because of the extremely flexible control of the vocal apparatus in humans as manifest in exploratory vocalization, humans can create new categories of sound not provided in their innate repertoire (e.g., cry or laughter). These new sounds can be graded, but they also develop into discrete syllable types by the second half year of life in canonical babbling (Oller, 1980; Stark, 1981). These discrete syllable types form the foundation for unlimited recombinations of syllables and thus the basis for an indefinitely large vocabulary. No NHP has ever been shown to develop such discrete, recombinable syllable types, although many songbirds seem to produce recombinable syllable-like elements.

Considerable research has been devoted to demonstrating that voluntary vocalization may be possible in NHP as manifest in vocal learning of, for example, either a very small number of sounds after extensive training/experience, or modifications in usage of innately

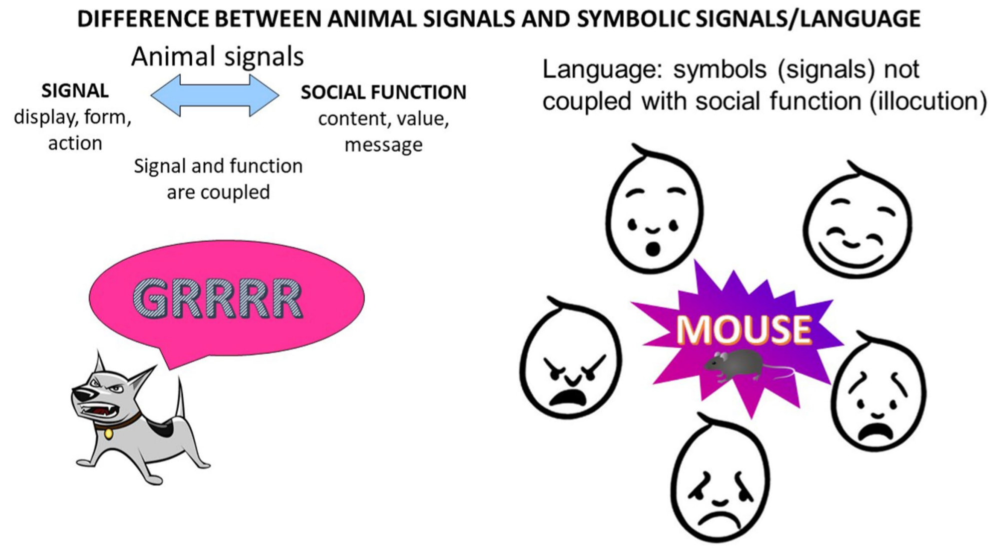


FIGURE 1
In vocal communication of nonhuman animals, each signal type ("call") tends to be coupled with a social function or class of functions, motivated by emotion (for example, growling motivated by Rage in dogs). In language, the signal ("mouse") and the illocutionary function, for example, expression of joy, play, fear, anger, or desire to share information, are not coupled. Linguistic signals can be used flexibly for any social function, that is, in the expression of any conceivable illocutionary force. And most importantly they can be produced purely as exploration, devoid of any social motive, proving their status of being decoupled from any immediate need and thus available for expression in any state and with any illocutionary purpose.

available sounds after similarly long periods of experience (Hopkins and Savage-Rumbaugh, 1986; Fischer, 2003; Hopkins et al., 2007; Taglialatela et al., 2012; Hage et al., 2016). But such research has not, to our knowledge, even attempted to quantitatively compare vocal learning in NHP with that which occurs even in a young human child, who can learn a large array of new sounds or sound sequences on single trials. The very weak vocal production learning in NHP contrasts sharply with the considerable vocal learning ability of humans and with the ability of NHP to learn how to interpret sounds they hear (Seyfarth and Cheney, 2010) or to produce non-vocal (non-phonatory) sounds imitatively (Hopkins et al., 2007).

A second difference, also presented in the list of four, is that humans interact vocally with positive affect, often with eye contact. Of course eye contact does occur communicatively in NHP (Bard et al., 2005; Kano et al., 2015; Heesen et al., 2021) although it appears usually to be brief and sometimes to be avoided (Kaplan and Rogers, 2002), perhaps as threatening. In contrast, a combination of directed gaze and non-aggressive vocalization exchange in human infancy is widely recognized as beginning in the first months of life and appears to produce bonding and a growing understanding in the infant of the possibility of sharing affect and recognizing the emotions and minds of others (Trevarthen, 1974, 1979; Terrace et al., 2022). This sort of sustained, affectively positive vocal exchange has never to our knowledge been observed in apes other than humans, although it may occur to some extent in singing birds (Fehér et al., 2017; Rivera-Cáceres et al., 2018) and in the New World callitrichid monkeys (Takahashi et al., 2013, 2015).

The third item on the list of four is vocal imitation of new sounds or sound sequences not previously occurring in the repertoire of the imitator. Humans learn to imitate particular syllable sequences that later become the most primitive words. No NHP has ever succeeded in producing more than an extremely limited variety of syllables, and that limited repertoire is uncertain (Kellogg and Kellogg, 1933; Gardner and Gardner, 1969; Gardner and Gardner, 1985). In general it appears NHP learn scarcely, if at all, to produce phonatory sounds not in their innate repertoires. Among the best examples is that Hopkins et al. (2007) claimed captive chimpanzees learned an "extended grunt" (with grunts already being in the innate repertoire), which the chimps used in attention getting with human caretakers. While interesting, this is hardly a demonstration of extensive phonatory learning. In contrast, many birds, e.g., blackbirds, corvids, and parrots have learned to imitate very large numbers of human syllables and more importantly very large repertoires of other sounds not in their innate repertoires.

Human infants by the second year learn to produce words that can be decoupled from any particular social function—each word production can occur with any one of a wide variety of functions. These words constitute semantic entities, that is, vocal symbols. No NHP has ever been shown to produce words in this way, although sign language learning and other visual domain learning in nonhuman apes has clearly produced cases of some non-vocal acquisition of word-like symbolic units with at least some functional flexibility (Fouts, 1987, 1991; Griebel et al., 2016). Many birds can imitate an extreme variety of sounds (Tchernichovski et al., 2001; Dalziell and Magrath, 2012; Lipkind et al., 2013). It appears the gray parrot Alex and his parrot colleagues could be the only *animals* ever to have been shown to produce imitative vocal words with illocutionary flexibility; still, even with Alex, the number of learned words was extremely small compared with word repertoires of even three-year-old humans (Pepperberg, 2010).

Because the list of four presents sequential steps of development in humans with each subsequent step being apparently dependent on the previous ones, it should be no surprise that apes other than humans have never reached a real word vocal symbolism stage, that is, step #4. Since they have never been shown to pass the first step, there is no logical basis to expect that they could pass the more advanced steps that are also required for language. Even though great apes have been shown to possess the cognitive requirements to learn gestural symbols when taught by humans, they have not independently evolved sign language either.

But why did humans, in some distant past, ever achieve the first step? This "ultimate" question (Tinbergen, 1963; Mayr, 1993) about the evolutionary conditions required for language foundations has rarely been considered explicitly. We have proposed there is solid logical reason to believe ancient hominins could not have gone further toward language without that first step (Oller et al., 2016). What evolutionary scenario could plausibly have presented circumstances where natural selection could have driven ancient hominins to produce sounds that were decoupled from any particular emotional state or social function? And why have animals such as song birds never gone on to language, even though many of them seem to have achieved the first step, and in some cases even step #2 or #3? Some bird species are able to imitate many sounds in their environment, whether they are made by other bird species or by human contraptions. Yet even in such cases, language has not emerged. One likely reason is a lower degree of sociality in birds than in humans, but there are other possibilities such as lesser brain complexity or size, the shorter time available for development, and the number of generations required for the evolutionary steps to occur.

Transition from emotional signal to symbol: the Fitness Signaling Theory

How was it possible for hominins to go from functionally fixed signals to functionally completely decoupled signals? Our answer has already been partially revealed, and we will suggest it also supplies a hint about how singing birds, who produce subsong (babble) prior to maturity (Ter Haar et al., 2021), may also have evolved to produce decoupled signals.

According to our proposal, the mechanisms of vocal control in ancient hominins had to be connected through natural selection to the Seeking System. We presume that the result was a much increased ability as well as an *inclination* on the part of ancient hominins to produce vocalizations that were decoupled from particular illocutionary functions and were often produced just for the pleasure of vocalizing. The decoupling presumably occurred first in infants, and those infants would have carried the decoupling capability into adulthood. Sounds thus produced through the Seeking System would have been possible to produce as endogenous exploration and in a wide variety of emotional/affective conditions both in infancy and later in life.

Why was this connection of the Seeking System to vocal production evolved in ancient hominins but not in other primates? We, along with Locke (2006, 2009) independently, have proposed that when ancient hominins became bipedal, and the hominin pelvis was necessarily changed (Washburn, 1960; Wells et al., 2012; Haeusler et al., 2021), selection pressure caused adaptations of development in the hominin infant in order to accommodate passage through the pelvis at birth (Bogin and Smith, 1996; Locke and Bogin, 2006). The combination of requirements is believed to have resulted in higher death rates in ancient hominin mothers, due to an "obstetrical dilemma" at the point of birth, and the problem persists; higher death rates clearly occur in modern human mothers at birth (Haeusler et al., 2021). More importantly for our reasoning, the accommodation is posited to have required a slowing of fetal development, resulting in a premature, altricial (helpless) newborn (Locke and Bogin, 2006; Hrdy, 2009). Furthermore, the slowing of development also yielded a longer period of relative helplessness, meaning hominin infants required (and modern human infants require) provisioning and protection for longer than their nonhuman ape relatives. That longer period was accompanied, in our reasoning, by heightened selection pressure on fitness signaling to elicit long-term investment from caregivers (Oller and Griebel, 2005, 2008; Locke, 2006, 2009).

Hominin infants thus competed against each other for care by vocalizing, since their helplessness greatly reduced their means of demonstrating wellness with physical movement (Oller and Griebel, 2014, 2021). Thus, we contend, the pressure to make vocalization more conspicuous and an increasingly more potent indicator of wellness resulted in increasing connection between the Seeking System and the vocal inclination of hominins. In this way our ancient ancestors came to use vocalization for the pleasure of Seeking and as a fitness signal, first in infancy as an advertisement that yielded caregiving, and later at all stages of maturation as an advertisement that yielded investment in the signaler in the form of mating, alliance formation, and cooperative activity. All this happened, according to our hypothesis, before language existed in the form of words and recombination of words in sentences.

This Fitness Signaling Theory (FST) is the only attempt we know of to explain the ultimate origin of exploratory vocalization in humans. The notion that exploratory vocalization could have been evolved directly *as a prerequisite for language* makes no sense, because natural selection cannot see into the future (Dawkins, 1996). If the capacity for exploratory vocalization had to evolve before language, there had to be an advantage to exploratory vocalization independent of language. The advantage of securing long-term caregiver investment through fitness signaling by altricial hominin infants suggests the solution.

We also contend that the value of fitness signaling provides a primary basis for maintaining massive vocal activity in human infants in modern times. We have plans to seek empirical support for the FST in the near future through experimentation, including monitoring physiological responses of caregivers and potential caregivers, who will listen to infant protophones in a variety of circumstances and to other kinds of sounds in the same circumstances.

Are there other reasons that vocalization evolved to become so important in human communication? Consider a property that makes vocalization particularly available for selection as a modality of communication. The vocal system in NHP is not often necessary for doing things beyond communicating (Oller and Griebel, 2021). The hands, for example, are used for manipulating objects, carrying things, climbing, and so on, and consequently hand gestures are often not available as communicative vehicles. But the vocal system, aside from its occasional functions in respiration (coughing, sneezing) and digestion (burping, hiccoughing) is available to be exploited as a signaling system because it is not required for many other functions. Perhaps all mammals have vocal signals partly because the vocal modality is almost always available to be targeted by natural selection for signaling. Thus the vocal capacity was surely open to extensive developments in ancient hominins.

Another property of vocalization that affords it an advantage over other kinds of possible fitness signaling is that receivers do not have to be looking in order to notice vocal signals. Vocal signals are effective even in the dark. These facts illustrate major advantages to vocal signaling over gestural or other visual forms of communication.

In addition, there were other conditions of life among ancient hominins favoring vocal fitness signaling in the infant. One was relatively large group sizes (Dunbar, 1996), affording protection from predators and reducing the pressure on silence to keep from alerting predators. Yet another factor was the growing tendency across hominin evolution for individuals to be cooperative breeders (Snowdon, 2004; Burkart et al., 2018), that is, many group members participated in caregiving for infants. As a result, there were many potential caregivers who might notice infant vocal fitness signals and invest more heavily in infants who produced especially effective signals.

Finally, we contend that there was an adaptation in hominins for conscious control of the glottis, owing to the same factor in hominin living that may have led to bipedalism, namely significant periods of waterside living (Hardy, 1960; Morgan, 1997; Tobias, 2002; Wrangham et al., 2009; Verhaegen et al., 2011; Joordens et al., 2019). Foraging for food by wading in water would have placed pressure on upright posture and bipedalism (Kuliukas et al., 2009), which is the apparent source of our altriciality at birth, and for the ability to consciously close off the glottis for foraging underwater in diving. The ability to consciously control glottal closure must have been accompanied by higher ability to control adduction of the vocal folds along with enhanced control of subglottal pressure required for phonation. Thus we reason that ancient hominins may have been more sensitive than other primates to natural selection pressures on fitness signaling by vocalization at the same point at which obligate bipedalism was making them more altricial.

The suggestion that diving may be not only associated with heightened conscious control of the glottis but with vocal learning in general seems plausible since many diving marine mammals are known to be vocal learners (Schusterman, 2008; Vernes et al., 2021). Of potentially similar interest is the fact that flight may also be associated with heightened conscious control of respiration and consequently with vocal learning capacities in thousands of bird species (Berg et al., 2019).

The Fitness Signaling Theory as a basis for vocal learning in humans is consistent with vocal learning in many species. The largest vocal repertoires, along with greatest variability and sources of vocal novelty are found regularly to be used for fitness advertising. Approximately 4,000 species of songbirds are believed to use their songs as signals of fitness to potential mates and as signals of their capability to protect their territories against invaders (Hausberger and Snowdon, 1997). So there is nothing that should be viewed as unusual in our proposing that human vocalization involves fitness signaling as perhaps its most fundamental motivation (Miller, 1996, 2000).

Additional features of vocal fitness signaling

Many types of vocalizations, not merely those motivated by the Seeking System, can serve as fitness signals. For example, mating songs produced by male birds are motivated by the Lust system as portrayed by Panksepp (1998), and are selected for by females for precision, variability, creativity, and perhaps for beauty to the ear of the beholder. Thus the form of courtship song is shaped by sensory biases of female birds. Birdsong is perhaps the most widely recognized type of vocal fitness signaling in the animal kingdom (Nottebohm, 1981; Baptista and Petrinovich, 1986; Hausberger and Snowdon, 1997; Kroodsma, 1999; Catchpole and Slater, 2003). There is good reason to think of much of birdsong and subsong in fledglings as being playful, and presumably motivated by the Seeking System. It seems clear that the Seeking System is involved in cases where males need to impress their audience with novel sound types (as in the case of humpback whales) and where males need to find new and exciting sounds in their environments to imitate so they can exceed the repertoires of their rivals. The Australian lyre bird can copy anything from a chain saw to a camera shutter sound with impressive accuracy (Dalziell and Magrath, 2012).

Territorial songs and choruses may be motivated by aggression and are demonstrations of vigor, stamina, and endurance to impress either single opponents in neighboring territories or to give competing groups living in nearby territories reason to stay away. In group-living social species, territorial songs may also reinforce group cohesion. Even sounds motivated by Rage can function as fitness signals in addition to their function in intimidating the individuals targeted by the aggressive act.

Securing caregiver investment may also be a general basis for selection of fitness signaling capabilities. Many animal infants (including most birds) face competition with siblings for caregiver investment through food and protection. This is not only true for siblings in the current litter/brood/clutch but also for consecutive single offspring births over the lifespan of the caregivers. Animals invest more in healthy offspring that in sickly ones, which they often abandon. Human infant mortality is still near 50% in some places in the world, where it is known that sickly infants are often abandoned. Even in the more prosperous modern world, neglect and abandonment are still focused on infants who fail to thrive (Locke, 2011).

Infant fitness advertisement has scarcely been investigated in other species, for example in birds, where song in adulthood is critically important, and where it has been reported on the basis of a broad survey that birds who sing in adulthood have a sort of bird "babbling" (subsong) in the fledgling stage (Ter Haar et al., 2021). Could subsong have been naturally selected long ago as fitness signaling before mature bird song existed? We know that the offspring of many animals beg for food vocally, and this has been attributed mostly to hunger, but we suspect that they may also be advertising their fitness (Rodríguez-Gironés et al., 1996). This question needs to be investigated.

Vocal creativity and high volubility might be useful for the caregivers of other species to determine offspring viability. Interestingly, the only other primates known to show "babbling" in the young are apparently the New World callitrichids (Elowson et al., 1998; Ghazanfar, 2013), including the marmosets and tamarins, which share the social system of cooperative breeding with humans. Perhaps cooperative breeding encourages fitness signaling in infant callitrichids. Nevertheless, these infants do not appear to use *novel* sounds in their "babbling." Rather they appear to use the sounds of the adult (inert) repertoire.

Human infants produce protophones from the first day of life and, if they are born prematurely, as soon as they can breathe independently (Oller et al., 2019). Protophones occur voluminously, much more frequently than any other vocalization type, thousands daily. But only a small proportion occur in interaction with caregivers (Long et al., 2020, 2022), revealing that protophones are predominantly motivated endogenously, presumably as seeking behaviors that result in parental investment even though infant protophones do not usually seem to be consciously noticed by parents. The rate of protophone production seems undiminished even when infants are alone in a room or no one is attending to them vocally (Delack, 1976; Iyer et al., 2016). According to our hypothesis, ancient hominin caregivers supplied the selection force on vocal fitness signaling in infants, but infants were capable of transmitting fitness information even when not directing the vast majority of their vocalizations to any caregiver. Ancient hominin infants broadcasted their fitness signals to anyone who *might be listening*. The same pattern of fitness signaling appears to be operative in the present, and modern human caregivers clearly show interest in the sounds their infants produce, trying to elicit them in face-to-face interaction and imitating infant sounds they have come to recognize (Stern, 1974; Trevarthen, 1979; Bornstein et al., 1992; Gratier and Devouche, 2011; Bornstein et al., 2015; Gratier et al., 2015).

Another line of reasoning that seems compatible with our proposal has been advanced by Levinson and colleagues, who have written of "cuteness selection" (Levinson, 2006a,b, 2022). They propose that human infants and ancient hominin infants may have used both protophones and other features of infancy to tap into caregiver tendencies to select infants based on emotional reactions to their lovability, which may have been, according to their reasoning, subject to runaway selection (Fisher, 1915). The idea of runaway selection incorporates integrally the notion that fitness signaling requires a real association between the signal that comes under selection pressure and wellness.

In many species, especially many species of birds, fitness signaling is seen to involve a learned and highly variable sound repertoire designed to impress potential sexual partners. It seems plausible that a learned and variable repertoire in infancy may also be used in birds, as in human infants, as fitness signals that elicit care.

Additional steps toward language after the emergence of exploratory vocalization

The claims of the Fitness Signaling Theory differ from those of other widespread attempts to account for the evolution of language, because our proposal aims to account for a very early step, a beginning without which subsequent steps toward language would not have been possible. A great deal of publication about the evolution of language (Chomsky, 1986; Bickerton, 1990; Pinker, 1994; Harnad, 1996; Deacon, 1997; Christiansen and Kirby, 2003; Niyogi, 2006; Chater et al., 2008; Berwick and Chomsky, 2016) addresses advanced features of language such as syntax and complex vocabulary, often without even a mention of the early adaptations we have proposed in the present paper, adaptations that seem to have broken hominins away from the primate background long before there was language.

But we need a more elaborate theory of how, after the evolution of vocal fitness signaling in early hominins, the next steps toward language could have been instantiated by natural selection pressures. The necessary pressures appear not to have applied to other species with massive vocal repertoires—otherwise many birds would surely have evolved language. So there must be adaptive advantages to evolving a wide variety of additional features necessary to language (beyond exploratory, fitness signaling vocalization) that occurred in hominins but no other vocal learning species.

Consider the conventionalization of syllables and syllable sequences to create words. Many suggestions have been made about advantages of evolving words. For example, words make possible the naming of group members in order to keep track of social contacts and alliances; cooperative hunting has been thought to require words to coordinate actions; naming objects has been thought to facilitate tool use and trading objects; and there have been many other suggestions as cited in Christiansen and Kirby (2003). There have also been intriguing suggestions about how prosociality of hominins and their presumably cooperative tendencies may have supported evolution in the direction of language (Tomasello, 1996, 2008; Tomasello et al., 2005; Nowak and Highfield, 2011; Kaplan, 2014, 2023). Note that all these suggestions appear to be dependent on *highly social conditions*. Indeed, the advantages of vocabulary and later syntax are obviously increased with increasing complexity of culture.

Hominin groups appear to have increased in size across evolution (Dunbar, 1993, 1996), and consequently their cultures must have been complex, with complex communicative needs. One special need in primates is grooming, and as hominin groups increased in size, Dunbar (1993), argued they may have come under pressure to use vocalization as a substitute for grooming, since friendly social vocalization could be transmitted to multiple individuals simultaneously. The argument reinstates an earlier notion of "grooming talking" in ancient hominins (Morris, 1967). After vocal fitness signaling was evolved, and flexible vocal repertoires were available, it appears increasing needs for more powerful communicative capabilities made ancient hominins sensitive to natural selection pressures that promoted the expansion of their vocal repertoires and the building of vocabulary and syntax from them, which would have served both fitness signaling and "grooming" needs.

It seems likely that one reason birds have not evolved language is that their lifestyles never had the range of social interrelations and consequent advantages of coordinated action that occurred in ancient hominins. Interestingly, the largest vocal repertoires in nonhuman animals do not coincide with a high degree of sociality. Some of the most elaborate known mating songs occur in species where potential mating partners see each other only once a year. For example, among marine mammals, mating songs have evolved in some solitary living baleen whales (Simon et al., 2010; Stafford et al., 2018) as well as in seals (Bjørgesæter et al., 2004) and walruses (Sjare et al., 2003), but not in the extremely socially-living dolphins. The same is true for the socially living parrots, which have elaborate social vocal repertoires. But their vocalizations in mating are not generally treated as "mating songs" although they are used to coordinate mating (Spoon, 2006). We suspect that very socially living animals do not require mating songs since potential mates know each other so well they can assess fitness of individuals based on long-term experience with them (personal communication, Drew Rendall). In contrast, if animals meet their potential mates only once a year, an elaborate song and/or dance may be necessary to provide the fitness information for mate selection.

A comment is called for regarding "signature whistles," vocalizations that appear to be indicators of the identity of individual bottle-nosed dolphins, which are highly social and strong vocal learners (Caldwell and Caldwell, 1965; Quick and Janik, 2012; Janik and Sayigh, 2013; Fearey et al., 2019). Each individual is claimed to invent a signature whistle by modification of other whistles. Signature whistles constitute around half the whistles produced by free-ranging dolphins and a much larger proportion in captive dolphins isolated from conspecifics. The whistles produced by dolphins in isolation have been interpreted as attempts to make contact with the prior group. Other members of the group are reported to use a slightly modified "copied" version of the individual's whistle perhaps to call the individual (Janik and Sayigh, 2013). The copying has been interpreted as "reference" to the individual and thus has been taken as a limited indication of semantics in wild dolphins. It should be noted that there have been empirical challenges to the very existence of signature whistles (McCowan and Reiss, 1995, 1997), but the claim of their existence remains a suggestion of semi-semantic evolution in nonhuman animals. There is also evidence that bottlenose dolphins can learn to associate other specific whistles as well as visual symbols with specific objects through operant conditioning (Herman and Forestell, 1985), an achievement that suggests parallels to the learning of visual symbols in chimpanzees, bonobos, and gorillas; see review in Tomasello (2017). Killer whales have family group repertoires of discrete whistles that partially overlap with those of closely related groups, but not with those of strangers (Ford, 1991). In social dolphins, all group members use all the whistles of the group as well. Thus for scientists to discriminate between usage as individual signature whistles or group marking repertoires, a whistle discrimination experiment in dolphins would be useful. If it could be shown that dolphins do use certain calls as signatures, this would be indeed a limited case for "naming" in the wild.

Apes other than humans are, like dolphins, extremely social and intelligent. But they appear never to have evolved a basis for complex vocal communication because, we contend, they never evolved a large vocal fitness signaling repertoire. Perhaps because they were less altricial at birth than hominins, because their group sizes were smaller, and because they showed less cooperative breeding, there was not sufficient pressure to evolve creative vocal fitness advertising. Even more important, as far as we know, no NHP evolved an extensive adaptation for voluntary control of the glottis.

Vocalizations integrated across a wide range of emotions

Through the connectivity of the Seeking System to vocal control, humans are motivated not only to explore sounds they essentially invent in vocal exploration, but also to explore the originally innate sounds associated with emotions such as Fear or Rage. Mature humans (especially actors) can toy with gradations between such sounds at will, can combine them in alternating patterns and use any of them in any emotional state. All of us can pretend to be crying or laughing at a chosen level of intensity (some more convincingly than others), and we can perform these vocal acrobatics even while we are talking.

Thus humans can add emotional flavoring in the form of prosodic contours or variations in pitch or amplitude when producing any kind of sound, including speech or otherwise innate signals such as shrieks, moans, or laughter. Humans also conventionalize various sounds drawn from the innate repertoire when, for example, saying "ha" (suggesting laughter) but invoking some special intended meaning, for example triumph, as in "I got you!." Or we can growl in rough and tumble play, or vocalize with pleasure during a massage, copulation, or in the anticipation of tickling. Everyday prosody often reflects the emotional state, the motivation and often the intended illocution of the sender in speech, a pattern that in some cases results in language-specific "pitch accents" (Pierrehumbert, 1979; Pierrehumbert and Hirschberg, 1990; Gussenhoven and Jacobs, 1998).

No other mammal appears to have such vocal flexibility. No other primate has a large repertoire of discrete syllable types nor an indefinitely large repertoire of syllable sequences, decoupled from any particular function. Instead, other primates have small repertoires of vocal types, each of which is graded to serve a relatively narrow class of possible functions. Human signals can also be graded, and the possibility of gradedness applies, not just to specifically emotional signals such as crying or laughter, but to every syllable, every word, every sentence of any language, all of which are, in accord with our proposal, possible to modulate through the Seeking System.

Conclusion

Summary

Much misunderstanding in the attempt to understand nonhuman animal and human communication has been caused by terminological missteps. Nonhuman animal signals are overwhelmingly about emotional states and illocutions, rather than constituting symbolic/semantic elements that must be detachable from emotional states and their accompanying illocutionary forces. Linguistic terms such as "reference" or "syntax" that have often been used in describing nonhuman animal vocalization are confusing rather than clarifying. Furthermore, research that restricts interpretation of vocal behavior to external observable actions occurring in particular situational contexts represents a failure to even address the primary goals of evolutionary science. We must develop understanding of the functions and motivations underlying vocal behavior if we are ever to develop a workable theory of the evolution of communication.

We have proposed that each call type is coupled to particular emotional/motivational states in nonhuman animals. Such vocal signals must be flexible enough to express, for example, gradations of intensity, to allow for audience effects, and to allow expressions that reflect mixed emotions. Humans also possess vocal signals, such as cry and laughter, that are commonly coupled to particular functions, and these signals have very similar properties to those of NHP. But even cry and laughter become very flexible in humans beyond early infancy. How did this occur in evolution?

We hypothesize that hominin vocal communication first diverged from the vocal communication of our competitor primates through the evolution of fitness signaling in primarily exploratory vocalizations. We have proposed that this divergence required a naturally-selected connection between the Seeking System proposed by Panksepp (1998) and the vocal control system of ancient hominins, making it possible for hominins to produce a wide variety of sounds that were decoupled from any particular emotional state or illocutionary intent. This decoupling allowed hominins to evolve further flexibility of vocalization, making possible learned vocalizations that could be used in any emotional state. The beginning of the break with the primate background appears to have occurred in the altricial hominin infant, who was selected to vocalize exploratorily and plentifully, thus maximizing the likelihood of long-term investment from caregivers, who noticed the vocal expression of well-being in the infant sounds.

Hominin infant development of vocal fitness signaling constituted the *first step* in producing a flexible learned and large vocal repertoire, according to the Fitness Signaling Theory. Subsequent steps were necessary because a large vocal repertoire does not by itself yield language. Additional steps were presumably naturally selected because of advantages of complex communication in the highly social, cooperatively-breeding hominins. A wide variety of social functions, such as group protection, hunting, foraging, tool use, vocal grooming, and trading were promoted by vocal signals that were possible to create once exploratory vocalization was established deeply enough to allow the evolution of words and sentences.

Looking forward

The Fitness Signaling Theory, largely in agreement with a similar proposal by Locke (2006), represents a departure from the predominant trend in research on primate communication, a trend that utilizes the misleading terminology critiqued in our article. The trend seems to apply blinders to its proponents by encouraging attention only to communicative similarities between humans and other animals to the practical exclusion of addressing the important differences. The approach suggests there is something unseemly about investigating human uniqueness, as if to do so would require us to go back to thinking in ways that were common two centuries ago. The current predominant trend thus becomes, in our view, hidebound, rejecting one of the most fundamental goals of biological research, which is to account, whenever possible, for species differences in terms of adaptation. Language is a massive adaptation, treated by some biological theorists as one of the major transitions since the origin of life (Maynard Smith and Szathmáry, 1997; Maynard Smith and Szathmáry, 2000; Szathmáry, 2015), and we are trying to account for the most fundamental adaptive changes that laid groundwork for the evolution of the whole range of language capabilities.

Consider an analogy: ancestral saurischians (reptile-hipped dinosaurs) did not fly, yet the only surviving descendants of dinosaurs are believed to be thousands of species of birds (Padian and Chiappe, 1998). Flying in vertebrates is a major adaptive change, worthy of major scientific attention. The current account suggests feathers were first evolved for functions such as thermoregulation. The reasoning that goes into this account invokes symmetry of feathers in the first saurischians that had them and descent by modification to yield additional adaptive steps toward feathers with asymmetrical features compatible with flight (Prum, 1999; Benton et al., 2019). We suggest language deserves a similar scientific effort, and the Fitness Signaling Theory represents a proposal for how adaptations necessary for language were first selected.

It is as if advocates of the predominant trend of research on primate communication deny that language is a major adaptation, because they seek to show that nonhuman primates possess all the fundamental features of language. Thus the advocates deny the importance of developing a strategy that might lead to evolutionary explanation as has occurred in evolutionary research on avian flight. The denial is not explicit but is instead implemented in a research strategy where attention is not focused on the nature of language as an adaptation that goes vastly beyond the communicative capabilities of NHP or any other animal.

A workable account of the human language adaptation requires recognition and detailed portrayal of the nature of the differences as well as the similarities between language and vocal communication in our closest relatives. We see hopeful signs in research on primate communication, because a few recent articles (Clay et al., 2015; Dezecache et al., 2020; Taylor et al., 2022, 2023) have begun to address the possibility of vocal functional (or illocutionary) flexibility in our ape relatives. We hope that beginning will soon lead to a more concerted effort to develop a truly comparative enterprise where the origins of language are assessed in a broader evolutionary perspective and through direct empirical studies of vocal communication of both our ape relatives and ourselves.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

UG and DO wrote the manuscript and conceived of the theoretical framework it expresses. All authors contributed to the article and approved the submitted version.

Funding

This work reported here was inspired by a Focus Group funded by the Konrad Lorenz Institute for Evolution and Cognition (KLI) Research in Klosterneuburg, Austria and conducted at the KLI in the summer of 2019. This article was also supported by grants from the National Institutes of Health, R01 DC015108 and R01 DC011027 to DO and the Plough Foundation Endowment which funds DO's Chair of Excellence at the University of Memphis.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adolphs, R., and Anderson, D. J. (2018). *The neuroscience of emotion: a new synthesis*. Princeton, NJ: Princeton University Press.

Anderson, D. J. (2022). *The nature of the beast: how emotions guide us*. New York, NY: Basic Books.

Anderson, D. J., and Adolphs, R. (2014). A framework for studying emotions across species. *Cell* 157, 187–200. doi: 10.1016/j.cell.2014.03.003

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

Baptista, L. F., and Petrinovich, L. (1986). Song development in the white-crowned sparrow: social factors and sex differences. *Anim. Behav.* 34, 1359–1371. doi: 10.1016/S0003-3472(86)80207-X

Bard, K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., and Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (*Pan troglodytes*). *Dev. Psychol.* 41, 616–624. doi: 10.1037/0012-1649.41.4.616

Benson, K. E. (2019). *Opioids*. Buffalo, NY: The Rosen Publishing Group, Inc.

Benton, M. J., Dhouailly, D., Jiang, B., and McNamara, M. (2019). The early origin of feathers. *Trends Ecol. Evol.* 34, 856–869. doi: 10.1016/j.tree.2019.04.018

Berg, K. S., Delgado, S., and Mata-Betancourt, A. (2019). Phylogenetic and kinematic constraints on avian flight signals. *Proc. R. Soc. B Biol. Sci.* 286:20191083. doi: 10.1098/rspb.2019.1083

Berwick, R. C., and Chomsky, N. (2016). *Why only us: language and evolution*, Cambridge, MA for Berwick MIT Press.

Bickerton, D. (1990). *Language and species*. Chicago, IL: University of Chicago Press.

Bjørgesæter, A., Ugland, K. I., and Bjørge, A. (2004). Geographic variation and acoustic structure of the underwater vocalization of harbor seal (*Phoca vitulina*) in Norway, Sweden and Scotland. *J. Acoust. Soc. Am.* 116, 2459–2468. doi: 10.1121/1.1782933

Bogin, B., and Smith, B. H. (1996). Evolution of the human life cycle. *Am. J. Hum. Biol.* 8, 703–716. doi: 10.1002/(SICI)1520-6300(1996)8:6<703::AID-AJHB2>3.0.CO;2-U

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., and Suwalsky, J. T. D. (2015). Mother-infant contingent vocalizations in 11 countries. *Psychol. Sci.* 26, 1272–1284. doi: 10.1177/0956797615586796

Bornstein, M. H., Tamis-LeMonda, C. S., Tal, J., Ludemann, P., Toda, S., Rahn, C. W., et al. (1992). Maternal responsiveness to infants in three societies: the United States, France, and Japan. *Child Dev.* 63, 808–821. doi: 10.2307/1131235

Burgdorf, J., and Panksepp, J. (2006). The neurobiology of positive emotions. *Neurosci. Biobehav. Rev.* 30, 173–187. doi: 10.1016/j.neubiorev.2005.06.001

Burkart, J., Martins, E., Miss, F., and Zürcher, Y. (2018). From sharing food to sharing information: cooperative breeding and language evolution. *Interact. Stud.* 19, 136–150. doi: 10.1075/bct.112.10bur

Byrne, B., Griebel, U., Wood, J. B., and Mather, J. A. (2003). Squid say it with the skin: a graphic model for skin displays in Caribbean reef squid (Sepiotheutis sepiodea). *Berl. Paläobiol. Abhangl.* 3, 29–35.

Caldwell, M. C., and Caldwell, D. K. (1965). Individualized whistle contours in bottle-nosed dolphins (*Tursiops truncatus*). *Nature* 207, 434–435. doi: 10.1038/207434a0

Call, J. (2008). "How apes use gestures: the issue of flexibility" in *Evolution of communicative flexibility: complexity, creativity and adaptability in human and animal communication*. eds. D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 235–252.

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

Catchpole, C. K., and Slater, P. J. B. (2003). *Bird song: biological themes and variations*, Cambridge, UK: Cambridge University Press.

Chater, N., Reali, F., and Christiansen, M. H. (2008). Restrictions on biological adaptation in language evolution. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1015–1020. doi: 10.1073/pnas.0807191106

Cheney, D. L., and Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Anim. Behav.* 36, 477–486. doi: 10.1016/S0003-3472(88)80018-6

Cheney, D. L., and Seyfarth, R. M. (1996). Function and intention in the calls of non-human primates. *Proc. Br. Acad.* 88, 59–76.

Chóliz, M., Fernández-Abascal, E. G., and Martínez-Sánchez, F. (2012). Infant crying: pattern of weeping, recognition of emotion and affective reactions in observers. *Span. J. Psychol.* 15, 978–988. doi: 10.5209/rev_SJOP.2012.v15.n3.39389

Chomsky, N. (1965). *Aspects of the theory of syntax*. New York, NY: MIT Press.

Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York, NY: Praeger.

Christiansen, M. H., and Kirby, S. (2003). Language evolution: consensus and controversies. *Trends Cogn. Sci.* 7, 300–307. doi: 10.1016/S1364-6613(03)00136-0

Clay, Z., Archbold, J., and Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *Peer J* 3:e1124. doi: 10.7717/peerj.1124

Clay, Z., Smith, C. L., and Blumstein, D. T. (2012). Food-associated vocalizations in mammals and birds: what do these calls really mean? *Anim. Behav.* 83, 323–330. doi: 10.1016/j.anbehav.2011.12.008

Clay, Z., and Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Anim. Behav.* 77, 1387–1396. doi: 10.1016/j.anbehav.2009.02.016

Conrad, B., and Schönle, P. (1979). Speech and respiration. *Arch. Psychiatr. Nervenkr.* 226, 251–268. doi: 10.1007/BF00342238

Crockford, C., and Boesch, C. (2003). Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. *Anim. Behav.* 66, 115–125. doi: 10.1006/anbe.2003.2166

Crockford, C., Gruber, T., and Zuberbühler, K. (2018). Chimpanzee quiet hoo variants differ according to context. *R. Soc. Open Sci.* 5:172066. doi: 10.1098/rsos.172066

Crockford, C., Wittig, R. M., Mundry, R., and Zuberbuhler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* 22, 142–146. doi: 10.1016/j.cub.2011.11.053

Dalziell, A. H., and Magrath, R. D. (2012). Fooling the experts: accurate vocal mimicry in the song of the superb lyrebird, Menura novaehollandiae. *Anim. Behav.* 83, 1401–1410. doi: 10.1016/j.anbehav.2012.03.009

Damasio, A. (1999). *The feeling of what happens: body and emotion in the making of consciousness*. New York, NY: Harcourt Brace and Co.

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

Davila Ross, M., Owren, M. J., and Zimmermann, E. (2010). The evolution of laughter in great apes and humans. *Commun. Integr. Biol.* 3, 191–194. doi: 10.4161/cib.3.2.10944

Dawkins, R. (1996). *Climbing mount improbable*. New York, NY: W. W. Norton and Col.

de Waal, F. (2011). What is an animal emotion? *Ann. N. Y. Acad. Sci.* 1224, 191–206. doi: 10.1111/j.1749-6632.2010.05912.x

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

de Waal, F. (2019). *Mama's last hug: animal emotions and what they tell us about ourselves*. New York NY: Norton.

Deacon, T. W. (1997). *The symbolic species*. New York, NY: W. W. Norton & Co. Ltd.

Delack, J. (1976). Aspects of infant speech development in 1st year of life. *Can. J. Linguist.* 21, 17–37. doi: 10.1017/S0008413100008033

Delack, J., and Fowlow, P. (1978). "The ontogenesis of differential vocalization: development of prosodic constrastivity during the first year of life" in *The development of communication*. eds. N. Waterson and C. Snow (New York, NY: Wiley), 93–110.

Dezecache, G., Zuberbühler, K., Davila-Ross, M., and Dahl, C. D. (2020). Flexibility in wild infant chimpanzee vocal behavior. *J. Lang. Evol.* 6, 37–53. doi: 10.1093/jole/lzaa009

Dunbar, R. (1993). Coevolution of neocortical size, group size, and language in humans. *Behav. Brain Sci.* 16, 681–694. doi: 10.1017/S0140525X00032325

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

Ekman, P. (1994). "All emotions are basic" in *The nature of emotion*. eds. P. Ekman and R. J. Davidson (New York, NY for Ekman: Oxford University Press), 15–19.

Elowson, A. M., Snowdon, C. T., and Lazaro-Perea, C. (1998). 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends Cogn. Sci.* 2, 31–37. doi: 10.1016/S1364-6613(97)01115-7

Evans, C. S. (1997). Referential signals. *Communication* 12, 99–143. doi: 10.1007/978-1-4899-1745-4_5

Evans, C. S., and Evans, L. (1999). Chicken food calls are functionally referential. *Anim. Behav.* 58, 307–319. doi: 10.1006/anbe.1999.1143

Evans, C. S., Evans, L., and Marler, P. (1993). On the meaning of alarm calls: functional reference in an avian vocal system. *Anim. Behav.* 46, 23–38. doi: 10.1006/anbe.1993.1158

Farmer, C. G. (2006). On the origin of avian air sacs. *Respir. Physiol. Neurobiol.* 154, 89–106. doi: 10.1016/j.resp.2006.04.014

Fearey, J., Elwen, S. H., James, B. S., and Gridley, T. (2019). Identification of potential signature whistles from free-ranging common dolphins (*Delphinus delphis*) in South Africa. *Anim. Cogn.* 22, 777–789. doi: 10.1007/s10071-019-01274-1

Fehér, O., Ljubičić, I., Suzuki, K., Okanoya, K., and Tchernichovski, O. (2017). Statistical learning in songbirds: from self-tutoring to song culture. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160053. doi: 10.1098/rstb.2016.0053

Fischer, J. (2003). "Developmental modifications in the vocal behavior of non-human primates" in *Primate audition: ethology and neurobiology*. A. A. Ghazanfar (New York, NY: CRC Press), 109–126.

Fisher, R. A. (1915). The evolution of sexual preference. *Eugen. Rev.* 7, 184–192.

Ford, J. K. B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.* 69, 1454–1483. doi: 10.1139/z91-206

Fouts, R. (1987). "Chimpanzee signing and emergent levels" in *Cognition, language and consciousness: integrative levels*. eds. G. Greenberg and E. Tobach, vol. *2* (Hillsdale, NJ: Lawrence Erlbaum), 57–84.

Fouts, R. (1991). *Next of kin: my conversations with chimpanzees*. New York, NY: Avon Books.

Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., and van Schaik, C. P. (2019). Multimodal communication and language origins: integrating gestures and vocalizations. *Biol. Rev.* 94, 1809–1829. doi: 10.1111/brv.12535

Fröhlich, M., Wittig, R. M., and Pika, S. (2016). Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. *R. Soc. Open Sci.* 3:160278. doi: 10.1098/rsos.160278

Furrer, R. D., and Manser, M. B. (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *Am. Nat.* 173, 400–410. doi: 10.1086/596541

Gardner, R. A., and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 165, 664–672. doi: 10.1126/science.165.3894.664

Gardner, B. T., and Gardner, R. A. (1985). Signs of intelligence in cross-fostered chimpanzees. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 308, 159–176. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2858873.

Ghazanfar, A. A. (2013). Multisensory vocal communication in primates and the evolution of rhythmic speech. *Behav. Ecol. Sociobiol.* 67, 1441–1448. doi: 10.1007/s00265-013-1491-z

Gibson, W. T., Gonzalez, C. R., Fernandez, C., Ramasamy, L., Tabachnik, T., Du, R. R., et al. (2015). Behavioral responses to a repetitive visual threat stimulus express a persistent state of defensive arousal in Drosophila. *Curr. Biol.* 25, 1401–1415. doi: 10.1016/j.cub.2015.03.058

Glanzer, M. (1958). Curiosity, exploratory drive, and stimulus satiation. *Psychol. Bull.* 55, 302–315. doi: 10.1037/h0044731

Gozoules, H., Gozoules, S., and Ashley, J. (1995). "Representational signaling in non-human primate vocal communication" in *Current topics in primate vocal communication*. eds. E. Zimmerman, J. D. Newman and U. Jürgens (New York, NY: Plenum Press).

Grant, S. G. N. (2015). The molecular evolution of the vertebrate behavioural repertoire. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371, 20150051–20150059. doi: 10.1098/rstb.2015.0051

Gratier, M., and Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Dev. Psychol.* 47, 67–76. doi: 10.1037/a0020722

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., and Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Front. Psychol.* 6:1167. doi: 10.3389/fpsyg.2015.01167

Green, J. A., Whitney, P. G., and 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. doi: 10.1037/a0024173

Griebel, U., and Oller, D. K. (2008). "Evolutionary forces favoring contextual flexibiliity" in *Evolution of communicative flexibility: complexity, creativity and adaptability in human and animal communication*. eds. D. K. Oller and U. Griebel (Cambridge MA: MIT Press), 9–40.

Griebel, U., Pepperberg, I. M., and Oller, D. K. (2016). Developmental plasticity and language: a comparative perspective. *Top. Cogn. Sci.* 8, 435–445. doi: 10.1111/tops.12200

Gussenhoven, C., and Jacobs, H. (1998). *Understanding phonology*. London UK: Arnold.

Haeusler, M., Grunstra, N. D. S., Martin, R. D., Krenn, V. A., Fornai, C., and Webb, N. M. (2021). The obstetrical dilemma hypothesis: there's life in the old dog yet. *Biol. Rev.* 96, 2031–2057. doi: 10.1111/brv.12744

Hage, S. R., Gavrilov, N., and Nieder, A. (2016). Developmental changes of cognitive vocal control in monkeys. *J. Exp. Biol.* 219, 1744–1749. doi: 10.1242/jeb.137653

Hamilton, W. D. (1963). The evolution of altruistic behavior. *Am. Nat.* 97, 354–356. doi: 10.1086/497114

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

Harnad, S. (1996). "The origins of words: a psychophysical hypothesis" in *Communicating meaning: the evolution and development of language*. eds. B. Velikovsky and D. Rumbaugh (Mahwah, NJ: Lawrence Erlbaum Assoc., Inc), 27–44.

Hausberger, M., and Snowdon, C. T. (Eds.) (1997). *Social influences on vocal development*. Cambridge, UK: Cambridge University Press.

Hauser, M. D., and Marler, P. (1993). Food associated calls in rhesus monkeys, I. Socioecological factors. *Behav. Ecol.* 4, 194–205. doi: 10.1093/beheco/4.3.194

Heesen, R., Austry, D. A., Upton, Z., and Clay, Z. (2022a). Flexible signalling strategies by victims mediate post-conflict interactions in bonobos. *Philos. Trans. R. Soc. B Biol. Sci.* 377:20210310. doi: 10.1098/rstb.2021.0310

Heesen, R., Bangerter, A., Zuberbühler, K., Iglesias, K., Neumann, C., Pajot, A., et al. (2021). Assessing joint commitment as a process in great apes. *iScience* 24:102872. doi: 10.1016/j.isci.2021.102872

Heesen, R., Sievers, C., Gruber, T., Clay, Z., and Studies, P. C. (2022b). "Primate communication: affective, intentional, or both?" in *Primate Cognitive Studies*. eds. B. L. Schwartz and M. J. Beran (Cambridge University Press) Available at: https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2021.0310

Heller, A. S., Fox, A. S., Wing, E. K., McQuisition, K. M., Vack, N. J., and Davidson, R. J. (2015). The Neurodynamics of affect in the laboratory predicts persistence of real-world emotional responses. *J. Neurosci.* 35, 10503–10509. doi: 10.1523/jneurosci.0569-15.2015

Helweg, D. A., Frankel, A. S., Mobley, J. R., and Herman, L. M. (1992). "Humpback whale song: our current understanding" in *Marine mammal sensory systems*. eds. J. A. Thomas, R. Kastelein and A. Y. Supin (Plenum), 459–483.

Herman, L. M., and Forestell, P. H. (1985). Reporting presence or absence of named objects by a language-trained dolphin. *Neurosci. Behav. Rev.* 9, 667–681. doi: 10.1016/0149-7634(85)90013-2

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cogn. Sci.* 30, 3–41. doi: 10.1207/s15516709cog0000_50

Hoefler, C. D. (2002). Is contest experience a trump card? The interaction of residency status, experience, and body size on fighting success in *Misumenoides formosipes* Araneae: Thomisidae. *J. Insect Behav.* 15, 779–790. doi: 10.1023/A:1021123523307

Hoemann, K., Gendron, M., and Barrett, L. F. (2017). Mixed emotions in the predictive brain. *Curr. Opin. Behav. Sci.* 15, 51–57. doi: 10.1016/j.cobeha.2017.05.013

Hollén, L. I., and Manser, M. B. (2007). Motivation before meaning: motivational information encoded in Meerkat alarm calls develops earlier than referential information. *Am. Nat.* 169, 758–767. doi: 10.1086/516719

Hopkins, W. D., and Savage-Rumbaugh, E. S. (1986). Vocal communication in the pygmy chimpanzee (*Pan paniscus*) as a result of differential rearing experiences. *Am. J. Primatol.* 10, 407–408.

Hopkins, W. D., Taglialatela, J. P., and Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Anim. Behav.* 73, 281–286. doi: 10.1016/j.anbehav.2006.08.004

Hrdy, S. B. (2009). *Mothers and others: the evolutionary origins of mutual understanding*. Cambridge, MA: Harvard University Press.

Ikemoto, S., and Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Rev.* 31, 6–41. doi: 10.1016/S0165-0173(99)00023-5

Ivanov, B., Hester, E. B., Martin, J. C., Silberman, W., Slone, A. R., Goatley-Soan, S., et al. (2020). Persistence of emotion in the process of inoculation: experiencing post-attack threat, fear, anger, happiness, sadness, and surprise. *Commun. Q.* 68, 560–582. doi: 10.1080/01463373.2020.1850492

Iverson, J. M. (2010). Multimodality in infancy: vocal-motor and speech-gesture coordinations in typical and atypical development. *Enfance* 3, 257–274. doi: 10.4074/S0013754510003046

Iyer, S. N., Denson, H., Lazar, N., and Oller, D. K. (2016). Volubility of the human infant: effects of parental interaction (or lack of it). *Clin. Linguis. Phonet.* 30, 470–488. doi: 10.3109/02699206.2016.1147082

Iyer, S. N., and Oller, D. K. (2008). Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss. *Volta Rev.* 108, 115–138. doi: 10.17955/TVR.108.2.603

Jablonka, E., Ginsburg, S., and Dor, D. (2012). The co-evolution of language and emotions. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 2152–2159. doi: 10.1098/rstb.2012.0117

Janik, V. M., and Sayigh, L. S. (2013). Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A.* 199, 479–489. doi: 10.1007/s00359-013-0817-7

Jhang, Y., and Oller, D. K. (2017). Emergence of functional flexibility in infant vocalizations of the first 3 months. *Front. Psychol.* 8:300. doi: 10.3389/fpsyg.2017.00300

Joordens, J. C. A., Feibel, C. S., Vonhof, H. B., Schulp, A. S., and Kroon, D. (2019). Relevance of the eastern African coastal forest for early hominin biogeography. *J. Hum. Evol.* 131, 176–202. doi: 10.1016/j.jhevol.2019.03.012

Jung, Y., Kennedy, A., Chiu, H., Mohammad, F., Claridge-Chang, A., and Anderson, D. J. (2020). Neurons that function within an integrator to promote a persistent behavioral state in Drosophila. *Neuron* 105, 322–333.e5. doi: 10.1016/j.neuron.2019.10.028

Jürgens, U. (1995). "Neuronal control of vocal production in non-human and human primates" in *Current topics in primate vocal communication*. eds. E. Zimmerman, J. D. Newman and U. Jürgens (New York, NY: Plenum Press).

Jürgens, U., and Ploog, D. W. (1976). Zur Evolution der Stimme. *Arch. Psychiatr. Nervenkr.* 222, 117–137. doi: 10.1007/BF02206613

Kano, F., Hirata, S., and Call, J. (2015). Social attention in the two species of Pan: bonobos make more eye contact than chimpanzees. *PLoS One* 10:e0129684. doi: 10.1371/journal.pone.0129684

Kaplan, G. (2008). Alarm calls and referentiality in Australian magpies: between midbrain and forebrain, can a case be made for complex cognition? *Brain Res. Bull.* 76, 253–263. doi: 10.1016/j.brainresbull.2008.02.006

Kaplan, G. (2014). Animal communication. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 661–677. doi: 10.1002/wcs.1321

Kaplan, G. (2023). Evolution of human language: duetting as part of prosociality and cognition. *Front. Ecol. Evol.* 11:1004384. doi: 10.3389/fevo.2023.1004384

Kaplan, G., and Rogers, L. J. (2002). Patterns of gazing in orangutans (*Pongo pygmaeus*). *Int. J. Primatol.* 23, 501–526. doi: 10.1023/A:1014913532057

Kellogg, W. N., and Kellogg, L. A. (1933). The ape and the child: a study of environmental influence upon early behavior. *Nature* 10, 421–422. doi: 10.1097/00006324-193311000-00009

Kent, R. D. (1998). *The speech sciences*. San Diego, CA: Singular.

Knight, C. (2016). Puzzles and mysteries in the origins of language. *Lang. Commun.* 50, 12–21. doi: 10.1016/j.langcom.2016.09.002

Krebs, J. R., and Dawkins, R. (1978). "Animal signals: mind reading and manipulation" in *Behavioural ecology: an evolutionary approach*. eds. J. R. Krebs and N. B. Davies (Sunderland, UK: Sinauer Associates), 380–402.

Kroodsma, D. E. (1999). "Making ecological sense of song development" in *The design of animal communication*. eds. M. D. Hauser and M. Konishi (Cambridge, MA: MIT Press), 319–342.

Kuliukas, A. V., Milne, N., and Fournier, P. (2009). The relative cost of bent-hip bent-knee walking is reduced in water. Homo: the journal of comparative. *Hum. Biol.* 60, 479–488. doi: 10.1016/j.jchb.2009.09.002

Kumar, V., Croxson, P. L., and Simonyan, K. (2016). Structural organization of the laryngeal motor cortical network and its implication for evolution of speech production. *J. Neurosci.* 36, 4170–4181. doi: 10.1523/jneurosci.3914-15.2016

Larsen, J. T., and McGraw, A. P. (2014). The case for mixed emotions. *Soc. Personal. Psychol. Compass* 8, 263–274. doi: 10.1111/spc3.12108

Launay, J., and Pearce, E. (2020). "Singing as an evolved behavior for social bonding: the ice-breaker effect, Beta-endorphins, and groups of more than 150 people" in *The Routledge companion to interdisciplinary studies in singing* (Milton Park, UK: Routledge), 136–145.

LeDoux, J. (1998). *The emotional brain*. London, UK: Weidenfeld and Nicolson.

Levinson, S. C. (2006a). Cognition at the heart of human interaction. *Disc. Stud.* 8, 85–93. doi: 10.1177/1461445606059557

Levinson, S. C. (2006b). "On the human "interaction engine"" in *Roots of human sociality: culture, cognition and interaction*. eds. N. J. Enfield and S. C. Levinson (Oxford, UK: Berg), 39–69.

Levinson, S. C. (2022). The interaction engine: cuteness selection and the evolution of the interactional base for language. *Philos. Trans. R. Soc. B Biol. Sci.* 377:20210108. doi: 10.1098/rstb.2021.0108

Liebal, K., and Oña, L. (2018). Mind the gap—moving beyond the dichotomy between intentional gestures and emotional facial and vocal signals of nonhuman primates. *Interact. Stud.* 19, 121–135. doi: 10.1075/is.17040.lie

Lipkind, D., Marcus Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M., and Tchernichovski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* 498, 104–108. doi: 10.1038/nature12173

Locke, J. L. (2006). Parental selection of vocal behavior: crying, cooing, babbling, and the evolution of language. *Hum. Nat.* 17, 155–168. doi: 10.1007/s12110-006-1015-x

Locke, J. L. (2009). Evolutionary developmental linguistics: naturalization of the faculty of language. *Lang. Sci.* 31, 33–59. doi: 10.1016/j.langsci.2007.09.008

Locke, J. L. (2011). *Duels and duets: why men and women talk so differently*. Cambridge, UK: Cambridge University Press.

Locke, J. L. (2021). The indexical voice: communication of personal states and traits in humans and other primates. *Front. Psychol.* 12:651108. doi: 10.3389/fpsyg.2021.651108

Locke, J. L., and Bogin, B. (2006). Language and life history: a new perspective on the evolution and development of linguistic communication. *Behav. Brain Sci.* 29, 259–280. doi: 10.1017/S0140525X0600906X

Long, H. L., Bowman, D. D., Yoo, H., Burkhardt-Reed, M. M., Bene, E. R., and Oller, D. K. (2020). Social and endogenous infant vocalizations. *PLoS One* 15:e0224956. doi: 10.1371/journal.pone.0224956

Long, H. L., Ramsay, G., Bene, E. R., Burkhardt-Reed, M. M., and Oller, D. K. (2022). Perspectives on the origin of language: infants vocalize most during independent vocal play but produce their most speech-like vocalizations during turn taking. *PLoS One* 17:e0279395. doi: 10.1371/journal.pone.0279395

Lorenz, K. (1951). Ausdrucksbewegungen höherer Tiere. *Naturwissenschaften* 38, 113–116. doi: 10.1007/BF00589766

Lorenz, K. (1971). *Studies in animal and human behaviors*. (R Martin, Trans.). Cambridge, UK: Harvard University Press.

Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E., et al. (2017). Social laughter triggers endogenous opioid release in humans. *J. Neurosci.* 37, 6125–6131. doi: 10.1523/JNEUROSCI.0688-16.2017

Marler, P., and Evans, C. (1996). Bird calls: just emotional displays or something more? *Ibis* 138, 26–33. doi: 10.1111/j.1474-919X.1996.tb04310.x

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

Maynard Smith, J., and Harper, D. (2003). *Animal Signals* Oxford, UK: Oxford University Press.

Maynard Smith, J., and Szathmáry, E. (1997). *The major transitions in evolution*. Oxford UK: OUP Oxford.

Maynard Smith, J., and Szathmáry, E. (2000). *The origins of life: from the birth of life to the origin of language*. Oxford, UK: OUP Oxford.

Mayr, E. (1993). Proximate and ultimate causations. *Biol. Philos.* 8, 93–94. doi: 10.1007/BF00868508

McCowan, B., and Reiss, D. (1995). Quantitative comparison of whistle repertoires from captive adult bottlenose dolphins (Delphinidae, *Tursiops truncatus*): a re-evalutation of the signature whistle hypothesis. *Ethology* 100, 194–209. doi: 10.1111/j.1439-0310.1995.tb00325.x

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

Messenger, J. B. (2001). Cephalopod chromatophores: neurobiology and natural history. *Biol. Rev.* 76, 473–528. doi: 10.1017/S1464793101005772

Miller, G. (1996). "Sexual selection in human evollution: review and prospects" in *Evolution and human behaviour: ideas, issues and applications*. eds. C. Crawford and D. Krebs (Mahwah, NJ: Lawrence Erlbaum).

Miller, G. (2000). *The mating mind: how sexual choice shaped the evolution of human nature*. New York, NY: Doubleday.

Montgomery, K. C. (1954). The role of the exploratory drive in learning. *J. Comp. Physiol. Psychol.* 47, 60–64. doi: 10.1037/h0054833

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

Morris, D. (1967). *The naked ape*. New York, NY: Dell.

Niyogi, P. (2006). *The computational nature of language learning and evolution*. Cambridge MA: MIT Press.

Nottebohm, F. (1981). A brain for all seasons: cyclical anatomical changes in song control nuclei of the canary brain. *Science* 214, 1368–1370. doi: 10.1126/science.7313697

Nowak, M. A., and Highfield, R. (2011). *SuperCooperators: altruism, evolution, and why we need each other to succeed*. New York, NY: Free Press.

Olds, J., and Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.* 47, 419–427. doi: 10.1037/h0058775

Oller, D. K. (1980). "The emergence of the sounds of speech in infancy" in *Child phonology, Vol 1: production*. eds. G. Yeni-Komshian, J. Kavanagh and C. Ferguson (New York, NY: Academic Press), 93–112.

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., and Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6318–6323. doi: 10.1073/pnas.300337110

Oller, D. K., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Lee, C.-C., et al. (2019). Preterm and full term infant vocalization and the origin of language. *Sci. Rep.* 9:14734. doi: 10.1038/s41598-019-51352-0

Oller, D. K., and Griebel, U. (2005). "Contextual freedom in human infant vocalization and the evolution of language" in *Evolutionary perspectives on human development*. eds. R. Burgess and K. MacDonald (Thousand Oaks, CA: Sage Publications), 135–166.

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

Oller, D. K., and Griebel, U. (2014). On quantitative comparative research in communication and language evolution. *Biol. Theory* 9, 296–308. doi: 10.1007/s13752-014-0186-7

Oller, D. K., and Griebel, U. (2021). Functionally flexible signaling and the origin of language. *Front. Psychol.* 11:626138. doi: 10.3389/fpsyg.2020.626138

Oller, D. K., Griebel, U., and Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Top. Cogn. Sci.* 8, 382–392. doi: 10.1111/tops.12198

Oller, D. K., Ramsay, G., Long, H. L., and Griebel, U. (2021). Protophones, the precursors to speech, dominate the human infant vocal landscape. *Philos. Trans. R. Soc. B Biol. Sci.* 376:20200255. doi: 10.1098/rstb.2020.0255

Oña, L. S., Sandler, W., and Liebal, K. (2019). A stepping stone to compositionality in chimpanzee communication. *PeerJ* 7:e7623. doi: 10.7717/peerj.7623

Oudeyer, P.-Y., and Smith, L. (2016). How evolution may work through curiosity-driven developmental process. *Top. Cogn. Sci.* 8, 492–502. doi: 10.1111/tops.12196

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

Owings, D. H., and Zeifman, D. M. (2004). "Human infant crying as an animal communication system" in *The evolution of communication systems: a comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge MA: MIT Press), 151–170.

Owren, M. J., Amoss, R. T., and Rendall, D. (2011). Two organizing principles of vocal production: implications for nonhuman and human primates. *Am. J. Primatol.* 73, 530–544. doi: 10.1002/ajp.20913

Owren, M. J., and Linker, C. D. (1995). "Some analysis methods that may be useful to acoustic primatologists" in *Current topics in primate vocal communication*. eds. E. Zimmerman, J. D. Newman and U. Jürgens (New York, NY: Plenum Press).

Owren, M. J., and Rendall, D. (2001). Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol. Anthropol.* 10, 58–71. doi: 10.1002/evan.1014

Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biol. Philos.* 25, 755–780. doi: 10.1007/s10539-010-9224-4

Padian, K., and Chiappe, L. M. (1998). The origin of birds and their flight. *Sci. Am.* 278, 38–47. doi: 10.1038/scientificamerican0298-38

Panksepp, J. (1998). *Affective neuroscience: the foundations of human and animal emotions (series in affective science)*. Oxford, UK: Oxford University Press.

Panksepp, J. (2000). The riddle of laughter: neuronal and psychoevolutionary underpinnings of joy. *Curr. Dir. Psychol. Sci.* 9, 183–186. doi: 10.1111/1467-8721.00090

Panksepp, J. (2005). Affective consciousness: core emotional feelings in animals and humans. *Consc. Cogn.* 14, 30–80. doi: 10.1016/j.concog.2004.10.004

Panksepp, J. (2010). Affective neuroscience of the emotional BrainMind: evolutionary perspectives and implications for understanding depression. *Dialogues Clin Neurosci*, 12, 533–545. doi: 10.31887/DCNS.2010.12.4/jpanksepp

Panksepp, J. (2011). Toward a cross-species neuroscientific understanding of the affective mind: do animals have emotional feelings? *Am. J. Primatol.* 73, 545–561. doi: 10.1002/ajp.20929

Panksepp, J., and Biven, L. (2012). *The archaeology of mind: neuroevolutionary origins of human emotion*. New York, NY: W. W Norton & Company.

Panksepp, J., Siviy, S., and Normansell, L. (1984). The psychobiology of play: theoretical and methodological perspectives. *Neurosci. Biobehav. Rev.* 8, 465–492. doi: 10.1016/0149-7634(84)90005-8

Papaeliou, C., Minadakis, G., and Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *J. Speech Lang. Hear. Res.* 45, 311–317. doi: 10.1044/1092-4388(2002/024)

Papini, M. R., Penagos-Corzo, J. C., and Pérez-Acosta, A. M. (2018). Avian emotions: comparative perspectives on fear and frustration. *Front. Psychol.* 9:2707. doi: 10.3389/fpsyg.2018.02707

Pepperberg, I. M. (2010). Vocal learning in Grey parrots: a brief review of perception, production, and cross-species comparisons. *Brain Lang.* 115, 81–91. doi: 10.1016/j.bandl.2009.11.002

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *J. Acoust. Soc. Am.* 66, 363–369. doi: 10.1121/1.383670

Pierrehumbert, J., and Hirschberg, J. (1990). "The meaning of intonational contours in the interpretation of discourse,"in *Intentions in communication*, Eds. P. R. Cohen, J. Morgan, and M. E. Pollack (Cambridge, MA: MIT Press). 271–311.

Pinker, S. (1994). *The language instinct*. *3rd* Edn. New York, NY: Harper Perennial.

Ploog, D. W. (1992). "The evolution of vocal communication" in *Nonverbal vocal communication*. eds. H. Papoušek, U. Jürgens and M. Papoušek (Cambridge, UK: Cambridge University Press), 6–30.

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Sci. Rep.* 5:13220. doi: 10.1038/srep13220

Prum, R. O. (1999). Development and evolutionary origin of feathers. *J. Exp. Zool.* 285, 291–306. doi: 10.1002/(SICI)1097-010X(19991215)285:4<291::AID-JEZ1>3.0.CO;2-9

Quick, N. J., and Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proc. R. Soc. B Biol. Sci.* 279, 2539–2545. doi: 10.1098/rspb.2011.2537

Rendall, D. (2021). Aping language: historical perspectives on the quest for semantics, syntax, and other rarefied properties of human language in the communication of Primates and other animals. *Front. Psychol*. 12:675172. doi: 10.3389/fpsyg.2021.675172

Rendall, D., and Owren, M. J. (2013). "Communication without meaning or information: abandoning language-based and informational constructs in animal communication theory" in *Animal communication theory: information and influence*. ed. U. E. Stegmann (Cambridge, UK: Cambridge University Press).

Rivera-Cáceres, K. D., Quirós-Guerrero, E., Araya-Salas, M., Templeton, C. N., and Searcy, W. A. (2018). Early development of vocal interaction rules in a duetting songbird. *R. Soc. Open Sci.* 5:171791. doi: 10.1098/rsos.171791

Rodríguez-Gironés, M. A., Cotton, P. A., and Kacelnik, A. (1996). The evolution of begging: signaling and sibling competition. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14637–14641. doi: 10.1073/pnas.93.25.14637

Rogers, L. J., Stewart, L., and Kaplan, G. (2018). Food calls in common marmosets, Callithrix jacchus, and evidence that one is functionally referential. *Animals* 8:99. doi: 10.3390/ani8070099

Scheider, L., Waller, B. M., Oña, L., Burrows, A. M., and Liebal, K. (2016). Social use of facial expressions in hylobatids. *PLoS One* 11:e0151733. doi: 10.1371/journal.pone.0151733

Scheiner, E., Hammerschmidt, K., Jürgens, U., and Zwirner, P. (2006). Vocal expression of emotions in normally hearing and hearing-impaired infants. *J. Voice* 20, 585–604. doi: 10.1016/j.jvoice.2005.09.001

Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., and Slocombe, K. E. (2013). Chimpanzee food calls are directed at specific individuals. *Anim. Behav.* 86, 955–965. doi: 10.1016/j.anbehav.2013.08.013

Schusterman, R. J. (2008). "Vocal learning in mammals with special emphasis on pinnipeds" in *The evolution of communicative flexibility: complexity, creativity, and adaptability in human and animal communication*. eds. D. K. Oller and U. Gribel (Cambridge, MA: Cambridge: MIT Press), 41–70.

Seyfarth, R. M., and Cheney, D. L. (2010). Production, usage, and comprehension in animal vocalizations. *Brain Lang.* 115, 92–100. doi: 10.1016/j.bandl.2009.10.003

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28, 1070–1094. doi: 10.1016/S0003-3472(80)80097-2

Shimada, Y. M. (2012). Infant vocalization when alone: possibility of early sound playing. *Int. J. Behav. Dev.* 36, 407–412. doi: 10.1177/0165025411431408

Simon, M., Stafford, K. M., Beedholm, K., Lee, C. M., and Madsen, P. T. (2010). Singing behavior of fin whales in the Davis Strait with implications for mating, migration and foraging. *J. Acoust. Soc. Am.* 128, 3200–3210. doi: 10.1121/1.3495946

Simonyan, K., Ackermann, H., Chang, E. F., and Greenlee, J. D. (2016). New developments in understanding the complexity of human speech production. *J. Neurosci. Off. J. Soc. Neurosci.* 36, 11440–11448. doi: 10.1523/JNEUROSCI.2424-16.2016

Sinha, C. (2004). "The evolution of language: from signals to symbols to system" in *The evolution of communication systems: a comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 217–235.

Sjare, B., Stirling, I., and Spencer, C. (2003). Structural variation in the songs of Atlantic walruses breeding in the Canadian high Arctic. *Aquat. Mamm.* 29, 297–318. doi: 10.1578/016754203101024121

Skinner, B. F. (1957). *Verbal behavior*. New York, NY: Appleton-Century-Crofts, Inc.

Snowdon, C. T. (2004). "Social processes in the evolution of complex cognition and communication" in *Evolution of communication systems: a comparative approach*. eds. D. K. Oller and U. Griebel (Cambridge, MA: MIT Press), 131–150.

Spoon, T. R. (2006). Parrot reproductive behavior, or who associates, who mates, and who cares. *Man. Parrot Behav.* 63:77. doi: 10.1002/9780470344651.ch8

Stafford, K. M., Lydersen, C., Wiig, Ø., and Kovacs, K. M. (2018). Extreme diversity in the songs of Spitsbergen's bowhead whales. *Biol. Lett.* 14:20180056. doi: 10.1098/rsbl.2018.0056

Stark, R. E. (1980). "Stages of speech development in the first year of life" in *Child phonology*. eds. G. Yeni-Komshian, J. Kavanagh and C. Ferguson, vol. *1* (New York, NY: Academic Press), 73–90.

Stark, R. E. (1981). Infant vocalization: a comprehensive view. *Infant Med. Health J.* 2, 118–128. doi: 10.1002/1097-0355(198122)2:2<118::AID-IMHJ2280020208>3.0.CO;2-5

Stegmann, U. E. (2013) in *Animal communication theory: information and influence*. ed. U. E. Stegmann (Cambridge, UK: Cambridge University Press).

Stern, D. N. (1974). "Mother and infant at play: the dyadic interaction involving facial, vocal, and gaze behaviors" in *The effect of the infant on its caregiver*. eds. M. Lewis and L. A. Rosenblum (New York, NY: Wiley), 187–213.

Stokoe, W. C. (1960). Sign language structure: an outline of the visual communication systems of the American deaf. *Stud. Linguist.* 8, 1–78.

Studdert-Kennedy, M. (2000). "Evolutionary implications of the particulate principle" in *The evolutionary emergence of language*. eds. C. Knight, M. Studdert-Kennedy and J. R. Hurford (Cambridge, UK: Cambridge University Press), 161–176.

Studdert-Kennedy, M. (2005). "How did language go discrete?" in *Language origins: perspectives on evolution*. ed. M. Tallerman (Oxford, UK: Oxford University Press), 48–67.

Suthers, R. A. (1999). "The motor basis of vocal performance in songbirds" in *The design of animal communication*. eds. M. D. Hauser and M. Konishi (Cambridge, MA: MIT Press), 37–62.

Sutton, D. (1979). "Mechanisms underlying learned vocal control in primates" in *Neurobiology of social communication in primates: an evolutionary perspective*. eds. H. D. Steklis and M. J. Raleigh (New York, NY: Academic Press), 45–67.

Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proc. Natl. Acad. Sci. USA* 112, 10104–10111. doi: 10.1073/pnas.1421398112

Taglialatela, J. P., Reamer, L., Schapiro, S. J., and Hopkins, W. D. (2012). Social learning of a communicative signal in captive chimpanzees. *Biol. Lett.* 8, 498–501. doi: 10.1098/rsbl.2012.0113

Taglialatela, J. P., Russell, J. L., Schaeffer, J. A., and Hopkins, W. D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS One* 6:e18852. doi: 10.1371/journal.pone.0018852

Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., et al. (2015). The developmental dynamics of marmoset monkey vocal production. *Science* 349, 734–738. doi: 10.1126/science.aab1058

Takahashi, D. Y., Narayanan, D. Z., and Ghazanfar, A. A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr. Biol.* 23, 2162–2168. doi: 10.1016/j.cub.2013.09.005

Taylor, D., Clay, Z., Dahl, C. D., Zuberbühler, K., Davila-Ross, M., and Dezecache, G. (2022). Vocal functional flexibility: what it is and why it matters. *Anim. Behav.* 186, 93–100. doi: 10.1016/j.anbehav.2022.01.015

Taylor, D., Gustafsson, E., Dezecache, G., and Davila-Ross, M. (2023). Vocal functional flexibility in the grunts of young chimpanzees. *iScience* 26:107791. doi: 10.1016/j.isci.2023.107791

Tchernichovski, O., Mitra, P. P., Lints, T., and Nottebohm, F. (2001). Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291, 2564–2569. doi: 10.1126/science.1058522

Ter Haar, S. M., Fernandez, A. A., Gratier, M., Knörnschild, M., Levelt, C., Moore, R. K., et al. (2021). Cross-species parallels in babbling: animals and algorithms. *Philos. Trans. R. Soc. B Biol. Sci.* 376:20200239. doi: 10.1098/rstb.2020.0239

Terrace, H. S., Bigelow, A. E., and Beebe, B. (2022). Intersubjectivity and the emergence of words. *Front. Psychol.* 13:693139. doi: 10.3389/fpsyg.2022.693139

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

Tinbergen, N. (1963). On aims and methods of ethology. *Z. Tierpsychol.* 20, 410–433. doi: 10.1111/j.1439-0310.1963.tb01161.x

Tobias, P. V. (2002). Some aspects of the multifaceted dependence of early humanity on water. *Nutr. Health* 16, 13–17. doi: 10.1177/026010600201600104

Tomasello, M. (1996). "The cultural roots of language" in *Communicating meaning: the evolution and development of language*. eds. B. Velikovsky and D. Rumbaugh (Mahwah, NJ: Lawrence Erlbaum Assoc., Inc), 275–308.

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

Tomasello, M. (2017). "What did we learn from the ape language studies?" in *Bonobos: unique in mind, brain, and behavior*. eds. B. Hare and A. Yamamoto (Oxford, UK: Oxford), 95–104.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/S0140525X05000129

Trevarthen, C. (1974). Conversations with a two-month-old. *New Sci.* 2, 230–235.

Trevarthen, C. (1979). "Communication and cooperation in early infancy. A description of primary intersubjectivity" in *Before speech: the beginnings of human communication*. ed. M. Bullowa (Cambridge, UK: Cambridge University Press), 321–347.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57. doi: 10.1086/406755

Tsien, J. Z., Chen, D. F., Gerber, D., Tom, C., Mercer, E. H., Anderson, D. J., et al. (1996). Subregion-and cell type–restricted gene knockout in mouse brain. *Cell* 87, 1317–1326. doi: 10.1016/S0092-8674(00)81826-7

Turner, J. H. (1996). The evolution of emotions in humans: a Darwinian–Durkheimian analysis. *J. Theory Soc. Behav.* 26, 1–33. doi: 10.1111/j.1468-5914.1996.tb00283.x

Verhaegen, M., Munro, S., Puech, P.-F., and Vaneechoutte, M. (2011). "Early hominoids: orthograde Aquarboreals in flooded forests?" in *Was man more aquatic in the past? Fifty years after Alister Hardy-waterside hypotheses of human evolution*. ed. M. Vaneechoutte, vol. *1* (Sharjah, United Arab Emirates: Bentham Science), 67–81.

Vernes, S. C., Janik, V. M., Fitch, W. T., and Slater, P. J. B. (2021). Vocal learning in animals and humans. *Philos. Trans. R. Soc. B Biol. Sci.* 376:20200234. doi: 10.1098/rstb.2020.0234

Washburn, S. L. (1960). Tools and human evolution. *Sci. Am.* 203, 63–75. doi: 10.1038/scientificamerican0960-62

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychol. Rev.* 20, 158–177. doi: 10.1037/h0074428

Wells, J. C. K., DeSilva, J. M., and Stock, J. T. (2012). The obstetric dilemma: an ancient game of Russian roulette, or a variable dilemma sensitive to ecology? *Am. J. Phys. Anthropol.* 149, 40–71. doi: 10.1002/ajpa.22160

Wrangham, R., Cheney, D., Seyfarth, R., and Sarmiento, E. (2009). Shallow-water habitats as sources of fallback foods for hominins. *Am. J. Phys. Anthropol.* 140, 630–642. doi: 10.1002/ajpa.21122

Young, L. (2012). *The chemistry between us: love, sex, and the science of attraction*. London, UK: Penguin.

Zahavi, A. (1975). Mate selection—a selection for a handicap. *J. Theor. Biol.* 53, 205–214. doi: 10.1016/0022-5193(75)90111-3

Zhu, F., Cizeron, M., Qiu, Z., Benavides-Piccione, R., Kopanitsa, M. V., Skene, N. G., et al. (2018). Architecture of the mouse brain synaptome. *Neuron* 99, 781–799.e10. doi: 10.1016/j.neuron.2018.07.007