



Review

Vocal functional flexibility: what it is and why it matters

Derry Taylor ^{a, b, *}, Zanna Clay ^c, Christoph D. Dahl ^{d, e}, Klaus Zuberbühler ^{b, g},
Marina Davila-Ross ^a, Guillaume Dezecache ^f



^a Department of Psychology, University of Portsmouth, Portsmouth, U.K.

^b Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

^c School of Psychology, Durham University, Durham, U.K.

^d Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan

^e Brain and Consciousness Research Center, Taipei Medical University Shuang-Ho Hospital, New Taipei City, Taiwan

^f Université Clermont Auvergne, CNRS, LAPSO, Clermont-Ferrand, France

^g School of Psychology and Neuroscience, University of St Andrews, St Andrews, U.K.

ARTICLE INFO

Article history:

Received 29 June 2021

Initial acceptance 11 October 2021

Final acceptance 8 December 2021

Available online 9 March 2022

MS. number: 21-00412R

Keywords:

communication

comparative

evolution

language

ontogeny

referentiality

Human speech is marked by a signal–function decoupling, the capacity to produce sounds that can fulfil a variety of functions, in contrast to nonverbal vocalizations such as laughter, cries and screams, which are functionally more rigid. It has been argued that this decoupling provides an essential foundation for the emergence of language, in both ontogeny and phylogeny. Although language has a deep evolutionary history, whether this capacity for vocal functional flexibility also exists in the vocal systems of nonhuman animals has been much overlooked. Reasons are multiple. Here, we propose to diagnose the problems that have thus far hindered progress on understanding the evolutionary basis of functional flexibility, an issue which can shed broader light on the evolution of language. In particular, we aim to clarify what vocal functional flexibility is, why it matters, why we believe it should be investigated in nonhuman animals and how this could be best achieved.

© 2022 The Author(s). Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The evolution of speech and language is a rich and vibrant research area. While a central debate in this field is the modality within which these communicative capacities originated (i.e. vocal versus gestural communication, see Pollick & de Waal, 2007), we do not take a position on this debate in this paper. Instead, we contend that regardless of the modality within which language originated and the modalities available to language users today, language is a primarily vocal form of communication in humans. Therefore, it is important to understand the phylogenetic precursors to human vocal communication to fully understand the origins of language. Many comparative approaches have involved looking for phylogenetic precursors of linguistic capabilities within the vocal communication system of nonhuman primates due to their close evolutionary relationship with humans. This includes evidence for vocal control (Boë et al., 2017; Fitch et al., 2016; Lameira & Shumaker, 2019), displacement in time between vocalization and the event the call pertains to (Lameira & Call, 2018), the ability to combine signals into longer sequences (Suzuki & Zuberbühler,

2019; Zuberbühler, 2018, 2020), reference to external events (Seyfarth et al., 1980; Seyfarth et al., 1980; Sievers & Gruber, 2016), intentional signal production (Byrne et al., 2017; Graham et al., 2020; Schel et al., 2013; Townsend et al., 2017) and adjustment to conspecifics' attentional and epistemic states (Crockford et al., 2012, 2017). Further, this includes the uncovering of basic laws that may apply to all animal communicative systems (Fedurek et al., 2017; Gustison et al., 2016; Heesen et al., 2019), although we note that some authors express a more sceptical view of the relevance of primate vocal behaviour to human language due to the underlying motivations to communicate in humans and other primates (Rendall & Owren, 2002).

By contrast, a series of works examining speech development in human infancy have encouraged comparative researchers to look at the phylogenetic distribution of prelinguistic vocal capacities that precede many of the general vocal capabilities mentioned above (Locke, 2009; Oller et al., 2016; Oller, Caskey, et al., 2019; Oller, Griebel, et al., 2019). A particularly important example is 'functional flexibility', or the capacity to produce sounds that are not tied to the fulfilment of one particular function but instead can be used in a multitude of ways to fulfil different functions on different

* Corresponding author.

E-mail address: derry.taylor@unine.ch (D. Taylor).

occasions. There are several documented cases of different call types being used in the same context across different populations and individuals of all great ape species. These could plausibly indicate some degree of flexibility in call function (bonobos, *Pan paniscus*: Hopkins & Savage-Rumbaugh, 1991; gorillas, *Gorilla gorilla*: Robbins et al., 2016; chimpanzees, *Pan troglodytes*: Watts, 2016; orang-utans, *Pongo abelii*: Wich et al., 2012). However, the variety of functional flexibility required for speech and language development depend on such flexibility within rather than between individuals (Oller et al., 2013). Further, many of the aforementioned ape sounds are unvoiced whereas infants' protophones (flexibly produced vocalizations including growls, squeaks and vocants, that is, vowel-like sounds) are voiced sounds (Oller, 2000), questioning their phylogenetic relevance. As precursors to the first fully formed canonical speech sounds, protophones are argued to provide the raw material for spoken language and are defined as nonreflexive voiced sounds with or without supraglottal articulation (Oller, 2000). Vocal functional flexibility has been suggested to be already evident in the protophone vocalizations (see Oller et al., 2013 supplementary movies for examples) of infants as young as 1 month old (Jhang & Oller, 2017). This may even be observed in preterm infants who regularly produce protophones without exogenous stimulation suggesting flexibility in vocal production (Oller, Caskey, et al., 2019). Oller and colleagues have suggested that vocal functional flexibility could constitute a phylogenetic point of divergence between human vocal communication and that of other animal species (Oller et al., 2013), which appear to depend on more functionally fixed vocal systems. For example, the phenomenon of 'functionally referential' calls, acoustically distinct call variants associated with specific external events and responses in conspecifics, has been routinely highlighted as a parallel between words and animal calls (Gifford et al., 2003; Kirchof & Hammerschmidt, 2006; Seyfarth et al., 1980; Slocombe & Zuberbühler, 2005; Townsend & Manser, 2013). However, compared to human speech, such calls would be classified as functionally fixed rather than functionally flexible.

Besides tentative lines of evidence from studies in bonobos and chimpanzees, the closest extant relatives of humans (Clay et al., 2015; Dezecache et al., 2021), there have been no other investigations, to our knowledge, of this capacity in nonhuman animal vocal communication. Oller, Griebel, et al. (2019) also investigated evidence for protophones, a functionally flexible signal type identified in human infants, in captive bonobos. The authors found that bonobo infants spontaneously produced vocalizations that were not laughs or screams and that showed rudimentary acoustic similarities with human infant protophones. By comparison, these calls were produced much less often than in human infants and were not used during vocal exchanges between infants and mothers as they were in humans. While the acoustic analysis was mainly descriptive, this study nevertheless suggests that nonhuman animals possess call types that are candidates for exhibiting functional flexibility. This supports our contention that searching for functional flexibility in nonhuman animal vocal communication is indeed worthwhile. Functional flexibility has been studied mostly in apes; other species not only of the primate order but beyond that order have been little studied.

The reasons for the apparent gap in understanding the evolutionary basis for functional flexibility in part stem from a strong historical bias in the literature towards functional reference in animal communication (Rendall, 2021). In this literature, researchers have tended to extensively search for evidence for signals that are tightly correlated with a particular referent within the external world. In the textbook example, vervet monkeys, *Chlorocebus pygerythrus*, produce acoustically distinct alarm calls depending on the predator type they observe (Seyfarth et al., 1980).

Indeed, in Macedonia and Evans's (1993) classical framework to study mammalian vocalizations (for a review see Townsend & Manser, 2013), animal calls only qualify as functionally referential, and thereby word-like, if there is a tight association between call type and its eliciting context. Nevertheless, given the importance of producing functionally flexible signals in human communication, understanding the evolutionary origins of this capacity seems to be an important endeavour that has, until now, been grossly overlooked. To stimulate further discussion and comparative research in this area, we here raise awareness of this issue. We explain what vocal functional flexibility is (and what it is not), why it matters and how it can be investigated in nonhuman species.

VOCAL FUNCTIONAL FLEXIBILITY: WHAT IT IS

Vocal functional flexibility is broadly defined as the capacity to produce sounds that can be used to fulfil any or multiple functions (Oller et al., 2013). As functionally flexible sounds do not have just one specific function, their function depends on how they are used at that given moment. Consider the following utterance (taken from Oller et al., 2013, p. 6318): 'the train has arrived'. Imagine this utterance is made in three circumstances, each corresponding to a particular illocutionary force (i.e. the act performed in producing a signal, such as greeting, warning or threatening, see Oller & Griebel, 2014). In one of these circumstances (a 'positive' one), the utterance is made by the speaker as she expresses great joy in seeing the train is arriving. In a second utterance (a 'negative' one), the behaviour of the speaker suggests a negative emotional state as the train is late. The utterance 'the train has arrived' comes to be produced in circumstances in which the speaker expresses her annoyance. One could imagine a third circumstance (a 'neutral' circumstance) over which the utterance is produced as a simple statement of the arrival of the train. Across all these utterances, the syntactic structure is equivalent and so is the lexical meaning (there is mention of a train arriving). By contrast, the illocutionary force and the perlocutionary effects (i.e. how receivers react to the signal) are different. In fact, the vocal unit 'the train has arrived' can in principle be produced to express anything (provided there is mutual acknowledgment in a given linguistic community), with any corresponding effect in receivers depending on the circumstances. It is a vocal unit marked by functional flexibility (i.e. the independence between the vocal unit and its function). Functionally flexible vocal units correspond to sounds that allow for a systematic variation in illocution and perlocution, whereby any illocutionary force (and the corresponding perlocutionary effect) can be produced.

When addressing the question of how early this capacity emerges in human ontogeny, it is obviously not easy to access the illocutionary force or perlocutionary effect of vocal units produced by nonverbal organisms, such as preverbal human infants. Therefore, one strategy has been to presuppose that infants produce vocalizations to communicate or express their affective state to caregivers and to stimulate reactions from them. In this regard, the illocutionary force of preverbal infant vocalizations could be assessed via the expression of their present affect. Such affective states refer to broad categories that do not necessarily specify more specific emotional states (i.e. fear), and may vary in associated arousal according to the dimension theory of affect (Weierich et al., 2010). Mothers have been shown to respond differently to the same infant protophone type on different occasions depending on the affective state expressed by the protophone on a particular occasion (Oller et al., 2013). Under this classification, one early form of evidence of vocal functional flexibility could be the decoupling between infants' affective state and their vocal production, and the systematic use of affect expression to elicit different responses on different occasions. We propose to call the decoupling between

vocal behaviour and affective state ‘affective decoupling’. There is already some evidence for this, as outlined below.

At birth, preterm infants produce a number of calls, some of them being qualified as cries and laughter, others as vocants, squeals and growls which are types of protophone (Oller et al., 2013; Oller, Caskey, et al., 2019). Interestingly, protophones are altogether five times more frequently produced than cries around birth time suggesting an important communicative function for these vocalizations (Oller, Caskey, et al., 2019; Oller, Griebel, et al., 2019). By 1 month of age infants produce cries mostly in situations that are deemed as negatively valenced for the infant, as evidenced by the infant’s facial patterns of behaviour. Laughter appears also to be affectively biased towards positive valence, only (or mostly) produced in situations deemed positive by human coders and interpreted as such by caregivers (Oller et al., 2013; Oller, Caskey, et al., 2019). In other words, laughter and cries by young human infants express fixed and mutually exclusive information: positive and negative states respectively. Correspondingly, infant laughs and cries systematically elicit different types of responses by caretakers. In particular, human caregivers respond with ‘encouraging’ behaviours in response to laughs and show behaviours that attempt to ‘change’ the situation in response to cries (Oller et al., 2013). These calls are each attached to a specific affective and physiological dimension and consistently produce a fixed type of behavioural response in receivers. However, by adulthood the acoustic characteristics of laughs and cries overlap with one another (Erickson et al., 2009) suggesting context may also be needed to disambiguate their meaning. By contrast, according to Oller and colleagues, infants produce protophones in various contexts to express positive, negative or neutral states. Correspondingly, these vocalizations systematically elicit different responses from caregivers depending on the affective state expressed. As such, unlike more functionally fixed calls like cries and laughter, infant protophones do not appear to be tied to a specific affective state or a singular illocutionary force. Importantly, positive, neutral and negative protophones are all reliably identified as protophones rather than cries or laughs (Jhang & Oller, 2017; Oller et al., 2013). Therefore, while there may be acoustic variation in protophones that depends on valence, at the level of the call type they are nevertheless a coherent yet functionally flexible call type that shows affective decoupling. Indeed, this is an important point, since human adults also produce acoustically different versions of the same word on different occasions depending on factors such as emotional state, but at the level of the overall word rather than acoustically distinct word variant, the word is still functionally flexible.

VOCAL FUNCTIONAL FLEXIBILITY: WHY IT MATTERS

To understand the evolutionary origins of language, comparative psychologists typically search for language-like features in the communication systems of nonhuman animals (Fitch, 2005; 2018; Fitch et al., 2005). Indeed, we have learnt a great deal about animal communication and language origins by adopting this approach. However, this approach may overlook other important processes, such as development (Eaton et al., 2018). Within a human lifetime, linguistic characteristics emerge directly from early communicative behaviours (Chater & Christiansen, 2010; Locke, 2009). Therefore, looking for mature linguistic capacities may be unhelpful if the more basic processes upon which these capacities rely are not present. Instead, looking at the initial ontogenetic stages of a complex capacity, such as language, can enable us to reconstruct its basis, understand how its components interact and establish how it could have evolved. For example, asking which phonemes of mature human language are shared with nonhuman animals is unlikely to lead

to any insightful comparisons because most nonhuman animals do not produce fully formed syllables (Engesser et al., 2015). However, comparing the extent to which the infrastructure (i.e. high levels of vocal motor control) required to develop such sounds is shared might be a more fruitful approach (Hockett & Hockett, 1960; Lameira et al., 2014; Oller, 2000). Therefore, we need to compare not just overall features of vocal communication in humans and other animals, but compare how those features develop (Griebel et al., 2016; Oller et al., 2016). In other areas related to biology, this kind of evolutionary–developmental approach has been widely adopted to elucidate the origins of complex phenotypes (Müller, 2007). Similarly, in psychology authors increasingly agree on the importance of combining comparative and developmental approaches (Bard & Leavens, 2014; Nielsen & Haun, 2016). Here, we also argue that this evo-devo perspective is also essential for understanding the evolutionary origins of language.

A central question to further elucidate the evolutionary origins of language, then, is what are the relevant features of developing communication systems that provide insights into the evolutionary origins? Several authors have reached this conclusion and have even attempted to identify these prelinguistic capacities in nonhuman animals. This includes articulatory homologies, that is, similarities in how sounds are articulated with the vocal apparatus between humans and great apes (Lameira et al., 2014), as well as babbling in pygmy marmosets, *Cebuella pygmaea* (Snowdon & Elowson, 2001) and in the subsong of avian species such as zebra finches, *Taeniopygia guttata* (Aronov et al., 2008). While acknowledging the importance of such phenomena in human vocal development, Oller et al. (2013) argued that a more foundational characteristic for the development of language is vocal functional flexibility. This, it is argued, provides a foundation upon which most other important linguistic capacities are built and which would not be possible if vocalizations were functionally fixed. Put simply, the argument is that if calls are functionally fixed, they will show little variability in form and function. However, most language-related vocal capacities presuppose flexibility in form and function in order for those capacities to emerge. In primatology it has also been argued that vocal flexibility plays an important role in language evolution. Primatologists have claimed, for example, that vocal flexibility is a capacity that likely emerged within the context of cooperative breeding which may favour the ability to flexibly negotiate social interactions (Lemasson et al., 2011; Zuberbühler, 2011).

By the third month of life, human infants have a vocal repertoire of functionally flexible sounds (protophones) and more functionally fixed sounds such as laughter and cries (Oller et al., 2013). Since protophones are the apparently only functionally flexible vocalizations in the human infant’s vocal repertoire and are also precursors to speech sounds (Oller et al., 1999), this supports the notion that vocal functional flexibility could provide a foundation for language to develop upon. Unlike functionally fixed vocal signals, which have necessarily limited functions, functionally flexible vocal signals can be adapted to be used for a range of purposes to provide a foundation for more elaborate patterns of communication to emerge. In classical ethology, animal calls, were considered ‘fixed signals’ naturally selected to fulfil a particular function (Seyfarth et al., 2010). Although, more recent studies on call types traditionally thought to be functionally fixed such as vervet monkey alarm calls (Seyfarth et al., 1980) have shown such calls are more flexibly produced across contexts than previously thought (see Price et al., 2015). Nevertheless, contextual flexibility is not tantamount to functional flexibility as discussed in the following section. However, there are very few studies designed to specifically test this, and a key next step is to examine whether nonhuman primates and other animals (particularly early in development) also show signs of vocal functional flexibility.

The study of vocal functional flexibility is also interesting in and of itself, independent of its relationship to language. Classical evolutionary animal communication theory is centred around the idea that a reliable relationship between signal form and function is required in order for signals to be able to evolve (Maynard Smith & Harper, 1995, 2003). From this perspective, functionally flexible calls are puzzling: how could a signal that is in some sense functionally unreliable evolve? It is possible that such pragmatic forms of communication are more efficient, but this may only be possible in a community of conspecifics with sufficiently sophisticated cognitive abilities for making inferences about a signaller's illocution. In such circumstances, a single call is needed to achieve a wide range of functions rather than a specialized call for each function. Indeed, the requisite cognitive abilities are widely known to be present in many animal species including both primates and birds (see Seyfarth & Cheney, 2015 for a review). Other benefits could include the ability to adapt communicative behaviours to local circumstances that may vary in both space and time which may be relevant to a wide range of species whose social environment and ecology are spatially and or temporally variable. Vocal functional flexibility thus matters not only for its ability to provide crucial new insights into language evolution, but also for its ability to expand the horizons of classical evolutionary animal communication theory.

HOW CAN THE EVOLUTION OF VOCAL FUNCTIONAL FLEXIBILITY BE INVESTIGATED?

To best situate the likely phylogenetic historical point of emergence of this capacity it is essential to search for this capacity in nonhuman animals. As noted previously, we are aware of two pieces of tentative research so far (Clay et al., 2015; Dezechache et al., 2021) both in great ape species (chimpanzees and bonobos). Our question is, why hasn't there been more research on this capacity in animals, considering the visibility of the original study in human infants (Oller et al., 2013)? In part, this may be related to the fact that even in the human infant literature, this topic has been studied systematically by a relatively small selection of researchers. In addition, animal communication research has traditionally focused on finding which contexts calls are mostly associated with in order to infer their meaning rather than contextual flexibility in vocal production (Maynard Smith & Harper, 1995, 2003). Indeed, it has been commonly assumed that animal calls are innate and inflexible (see Nieder & Mooney, 2020). Here, we envision three primary reasons that may have hindered progress on this question: (1) a lack of conceptual clarity in the framework and/or misconceptions from the animal behaviour community; (2) a lack of a clear protocol to be used in comparative research; (3) pending conceptual issues with the framework. We review below these strands of reasons and hope they will then be clarified.

Uncertainty and Misconceptions

One frequent reaction to the framework of vocal functional flexibility is that it is a trivial phenomenon because many animals are already known to produce single vocal units in a great diversity of contexts. For example, lilac-crowned amazons, *Amazona finschi*, produce song notes across a wide variety of behavioural contexts (Montes-Medina et al., 2016). However, although the two concepts (i.e. contextual flexibility and functional flexibility) are understandably confusable, the production of certain vocal units in a variety of contexts is distinct from vocal functional flexibility. Functional flexibility is the notion that certain vocal units can express different information on different occasions and thereby fulfil

a variety of functions. Although vocal functional flexibility does require vocal contextual flexibility, they are not equivalent. For instance, Guinea baboons, *Papio papio*, produce grunts in a diversity of contexts (Maciej et al., 2013) including foraging and travelling but also in affiliative, infant handling and greeting situations. However, whether these calls are expressed differently on different occasions (i.e. affective decoupling) and achieve different functions systematically depending on how they are expressed, has not yet been shown. To determine whether this is the case, research would be required to disambiguate whether Guinea baboons' grunts are expressed differently on different occasions (i.e. with different accompanying communicative cues) and whether they systematically fulfil different functions depending on how they are expressed. An important alternative is that the same call may perform the same function in different contexts. For example, a call may refer to the presence of a predator, but the actual referent may differ between contexts, with context providing the necessary information for listeners to infer the identity of the predator (Scarantino & Clay, 2015).

Within animal vocal communication, adult baboon grunts (Maciej et al., 2013; Rendall et al., 1999), adult vervet monkey alarm calls (Price et al., 2015), adult Japanese macaque, *Macaca fuscata yakui*, coo calls (Koda, 2004), adult and juvenile mangabey, *Cercocebus torquatus atys*, grunts (Range & Fischer, 2004), adult bonobo high hoots (Schamberg et al., 2016), adult chimpanzee grunts (Slocombe & Zuberbühler, 2010), close calls in banded mongooses, *Mungos mungo* (Jansen et al., 2012), rumble, snort and trumpet calls in Asian elephants, *Elephas maximus* (Stoeger & Baotic, 2021) and domestic dog, *Canis familiaris*, growls (Taylor et al., 2009), are all produced in a variety of contexts that have no apparent link with each other. This is indicative of vocal contextual flexibility. Such patterns are also not limited to these species and are also present in many other phylogenetic groups including squirrels, meerkats and a range of bird species including warblers and sparrows (see Seyfarth et al., 2010 for a review). However, it might well be that a 'supra' context (that would encompass several contexts) can explain why a given vocal unit is used across contexts (e.g. vervet monkeys' terrestrial alarm calls may all be given in the supra context 'threat'), with the contextually flexible vocal unit having a single particular function (e.g. warning conspecifics). There surely are examples in which no 'supra context' can be found, a situation that may allow for speculation (but no firm conclusion) that the vocal unit in question qualifies as functionally flexible. In fact, rationalizing at the level of the 'context' of vocal production cannot give us information about the way in which a vocal unit is used, and, a fortiori, cannot tell us whether a given vocal unit can fulfil a variety of functions on different occasions. Based on the range of species that appear to show contextually flexibly vocal communication, we are inclined to believe vocal functional flexibility is a widespread phenomenon in animal communication, but this is a mere hypothesis that awaits empirical confirmation.

A Protocol to Examine Vocal Functional Flexibility in Nonhuman Animals

How are we to determine an animal's illocutionary force when they produce vocal signals? How can we recognize whether they induce consistent behavioural effects in receivers? The original study (i.e. Oller et al., 2013) suffered less from these problems: infants' illocutionary force could be determined by their likely affective state upon vocalizing as seen by their facial behaviour. This could be judged by human coders, who are (or at least are biologically prepared to be) caregivers in their lifetime. Similarly, the effects on receivers could be appropriately determined by human

coders (i.e. whether the caregiver consoled the infant, played with them, etc.). In nonhuman animals, Oller et al. (2013) proposed using facial action coding systems to delve into the affective states animals are in when they produce calls. Facial action coding systems have been developed for a number of mammal species (Caeiro et al., 2013, 2017; Julle-Daniere et al., 2015; Parr et al., 2007; Wathan et al., 2015). This solution might not be practical, however, as it requires understanding the muscular configuration associated with various affective states in nonhuman animals (Mielke et al., 2020). Also, facial configuration may require a fully mature muscular system, which hampers its use in young animals. As it stands, general facial configurations (rather than a specific configuration of muscular action units) known to be specifically used in positive contexts (e.g. the open-mouth face of play in a variety of mammal species including primates, dogs, foxes, bears and hyaenas, see Palagi et al., 2016) or negative ones (e.g. the bared-teeth displays in many mammal species including canines and primates, see Rogers & Kaplan, 2002) may be used. Since facial expressions may not always reliably indicate affective state (Fridlund & Russell, 2006), we also recommend using other signals of affect, such as piloerection and postural behaviour and a focus on circumstances that are clearly appetitive or aversive for the animal. In captivity, other markers may be used insofar as they can help delve into the valence of the state animals are in when producing calls. For example, a growing body of research is adopting noninvasive techniques for identifying physiological markers of affective state such as infrared thermal imaging (e.g. Ioannou et al., 2015).

A more comprehensive approach to infer contextual information from interactive animal behaviour can be achieved by means of behavioural classifiers that single out the individuals, identify and track them over time, estimate their poses and classify their behaviours automatically from video footage. One such approach has recently been implemented using deep learning (Marks et al., 2020). In further steps, this approach would allow unsupervised classification of interactive animal behaviour, avoiding any biases towards a predicted behavioural repertoire. Hence, to determine the relevant units of functional flexibility, while filtering out the stochastic ones, one can conduct an algorithmic search for repeated co-occurrences of a vocal signal (sender) and sequences of behavioural units exhibited by the sender (illocutionary force) and by the recipient (perlocutionary effect).

Regarding receiver responses, a focus on infant–mother dyads (besides its immediate relevance to the original study in humans) may help to clarify the consistency between infants' likely affective state/s (as a proxy for illocutionary force) and maternal reaction. Herein, the mother could respond proactively (taking part in the activity of the infant to continue it), protectively (intervening to withdraw the infant from an activity or withdraw from her own activity with the infant), empathically (responding by soothing the infant or containing its distress) or neutrally (no intervention). In natural circumstances and in certain species, it may happen that mothers do not always react strongly to the infant's vocal production, especially at certain time points in development, such as during the weaning conflict phase (Weary et al., 2008). For these cases, or among species that spend less time with mothers early in development due to species-typical sociality, it may be useful to look at variation in behavioural responses to the same call type on different occasions from more distanced kin or nonkin. Importantly, such individuals may be more motivated to respond to a subject's calls, making it more likely that if functionally flexibility is present, it will be revealed by observations.

Finally, as functionally flexible vocal units can be used to fulfil a variety of functions, there may be individual variation in how and when vocalizations are produced to elicit different responses

in receivers. It might indeed be that one infant chimpanzee primarily uses grunts to attract its mother in negative situations, while another primarily uses it in play contexts. Interindividual variability in call functional use may then provide crucial insights into vocal functional flexibility. In natural circumstances, it is difficult to establish whether a given signal is more often used with one particular function because this might simply reflect the distribution of situations to which the animal is exposed to. However, variation in how a call is used within rather than between subjects provides the most convincing evidence of functional flexibility.

Pending Conceptual Issues

Although utterances of functionally flexible vocal units are acoustically similar enough to one another to be grouped together, such sounds may still be perceived differently by receivers. For example, human infant protophones produced in positive, neutral and negative circumstances may in fact constitute three distinct vocal units rather than a single call type being used in different ways. When considering the example 'the train is arriving', it appears that this sequence of words under various illocutionary forces would still have more in common than with a minimally different sequence of words (e.g. 'the plane is arriving'). This is not solely by virtue of a similar syntactic structure or lexical meaning, but also by virtue of similar acoustical features and acoustical constraints, regardless of what is perceived by receivers. Similarly, previous studies into functional flexibility in chimpanzees (Dezecache et al., 2021) and bonobos (Clay et al., 2015) found systematic acoustic variation within call types that flexibly expressed different affective states on different occasions. Importantly, a recent study on infant chimpanzee call acoustics (Dezecache et al., 2021) showed that the call types were spontaneously grouped into distinct categories by machine learning (see also Taylor et al., 2021). This suggests that despite systematic acoustic variation within call types, functionally flexible call types in the infant chimpanzee vocal repertoire nevertheless represent a single coherent unit of communication. In fact, responses from receivers are expected to be consistent with the illocutionary force of the signaller, which may correspond to changes in motivational and/or emotional state that are encoded into call acoustics as predicted by Morton's (1977) motivational structural rules. Indeed, Morton's (1977) predictions have been empirically shown to predict vocal structure between call types in humans and nonhuman animals (see Briefer, 2012 for a full review). A focus on broad vocal categories, which may sound (and appear) overall more similar to one another than to other calls of the repertoire, may therefore be the basis for the delineation of vocal units within a repertoire, and their subsequent qualification as functionally flexible or not. Indeed, analyses of human infant protophones have been conducted at the level of the call type, rather than acoustically distinct protophone variants that relate to affective state (Jhang & Oller, 2017; Oller et al., 2013). Acoustic analyses can support this decision. For example, if an apparently functionally flexible vocal unit shows significantly different acoustic properties on different occasions depending on the illocutionary force, cluster analyses and distance metrics can be used to ascertain the extent to which these calls are similar to or different from each other and other call types (see Taylor et al., 2021). In particular, the use of 'fuzzy' cluster analysis may be useful, as this technique is able to quantify acoustic gradedness within a vocal repertoire and has been recently successfully applied in this way among nonhuman primates (e.g. Wadewitz et al., 2015; Taylor et al., 2021). Additionally, similarities and differences in the kinds of behavioural responses that a call elicits depending on the illocutionary force

might be informative. In some cases, this can be assessed through playback experiments, although note that this will not be feasible for all species and call types, especially for those whose communication occurs largely in a dyadic context.

Summary

Overall, we suggest three key challenges in the exploration of vocal functional flexibility in nonhuman animals. First, vocal functional flexibility has suffered from conceptual unclarity and misconceptions in the animal communication literature. In particular, it is important to acknowledge that functional flexibility does not correspond to contextual flexibility, but instead refers to when the effect of a communicative act depends on what the signaller does with the signal (i.e. the illocutionary force), which is typically achieved through accompanying behavioural cues produced by the communicator. Second, there are limited methodological guidelines on how to evaluate vocal functional flexibility in nonhuman subjects. We suggest the use of reliable behavioural indicators of affect, possibly supported by automated classification, could help in this regard. On the receiver side it may be beneficial to focus on receivers with a motivation to respond to subjects' calls, such as mothers and kin. Finally, there are outstanding conceptual issues, such as how to interpret systematic acoustic variation in calls that may in turn have implications for whether functional flexibility is present or not. We advocate the use of cluster analyses and distance metrics to show the relationships between call types within a repertoire and evaluating whether seemingly functionally flexible units are coherent communicative units.

CONCLUSION

In this article, our aim was to invite our colleagues working in the field of animal communication, and specifically those interested in comparative approaches to language evolution, to systematically look for signs of functional flexibility in the vocal repertoire of their species of interest. This capacity, we believe, is foundational to speech in humans as it is one of the earliest speech-related capacities found in ontogeny. In addition, it is an interesting phenomenon on its own terms, raising novel questions for classical animal communication theory to contend with.

Importantly, we clarify here how and why vocal functional flexibility does not equate with vocal contextual flexibility, the latter being a fairly widespread phenomenon in animal communication. Vocal functional flexibility requires looking at transient states that relate to changes in functionality and that may reveal flexibility in the functional use of particular vocal units on different communicative occasions. Although this is particularly challenging in nonhuman animals, we believe we urgently need to develop ways in which this can be approached. Such investigation will inevitably be fruitful, as it will lead to a better understanding of the phylogenetic distribution of vocal functionally flexibility, a capacity that is undeniably central to speech as we know it.

Author Contributions

All authors contributed to the conceptualization of the article. D.T. and G.D. wrote the original draft. All authors contributed towards review and editing.

Declaration of Interest

There are no conflicts of interest between the authors of this paper.

Funding

The primary funder of this research was the 'University of Portsmouth, Department of Psychology' and the grant ID is 'PhD Bursary' which was awarded to 'Derry Taylor'.

Acknowledgments

We thank Kimborough Oller and Ulrike Griebel for valuable discussions on the topic which were instrumental in the development of this paper. We thank the Editor and an anonymous referee for their valuable comments on the manuscript.

References

- Aronov, D., Andalman, A. S., & Fee, M. S. (2008). A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science*, 320(5876), 630–634. <https://doi.org/10.1126/science.1155140>
- Bard, K. A., & Leavens, D. A. (2014). The importance of development for comparative primatology. *Annual Review of Anthropology*, 43(1), 183–200. <https://doi.org/10.1146/annurev-anthro-102313-030223>
- Boë, L. J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., Becker, Y., Rey, A., & Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS One*, 12(1), Article e0169321. <https://doi.org/10.1371/journal.pone.0169321>
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, 288(1), 1–20. <https://doi.org/10.1111/j.1469-7998.2012.00920.x>
- Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition*, 20(4), 755–769. <https://doi.org/10.1007/s10071-017-1096-4>
- Caeiro, C. C., Burrows, A. M., & Waller, B. M. (2017). Development and application of Cat-FACS: Are human cat adopters influenced by cat facial expressions? *Applied Animal Behaviour Science*, 189, 66–78. <https://doi.org/10.1016/j.applanim.2017.01.005>
- Caeiro, C. C., Waller, B. M., Zimmermann, E., Burrows, A. M., & Davila-Ross, M. (2013). OrangFACS: A muscle-based facial movement coding system for orangutans (*Pongo* spp.). *International Journal of Primatology*, 34(1), 115–129. <https://doi.org/10.1007/s10764-012-9652-x>
- Chater, N., & Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science*, 34(7), 1131–1157. <https://doi.org/10.1111/j.1551-6709.2009.01049.x>
- Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, 3, e1124. <https://doi.org/10.7717/peerj.1124>
- Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, 22(2), 142–146. <https://doi.org/10.1016/j.cub.2011.11.053>
- Crockford, C., Wittig, R. M., & Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances*, 3(11), Article e1701742. <https://doi.org/10.1126/sciadv.1701742>
- Dezechache, G., Zuberbühler, K., Davila-Ross, M., & Dahl, C. D. (2021). Flexibility in wild infant chimpanzee vocal behavior. *Journal of Language Evolution*, 6(1), 37–53. <https://doi.org/10.1093/jole/lzaa009>
- Eaton, T., Hutton, R., Leete, J., Lieb, J., Robeson, A., & Vonk, J. (2018). Bottoms-up! Rejecting Top-down human-centered approaches in comparative psychology. *International Journal of Comparative Psychology*, 31(1), 1–19. <https://doi.org/10.5070/P4311037589>
- Engesser, S., Crane, J. M., Savage, J. L., Russell, A. F., & Townsend, S. W. (2015). Experimental evidence for phonemic contrasts in a nonhuman vocal system. *PLoS Biology*, 13(6), Article e1002171. <https://doi.org/10.1371/journal.pbio.1002171>
- Erickson, D., Menezes, C., & Sakakibara, K. (2009). Are you laughing, smiling or crying?. In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference* (pp. 529–537). <https://doi.org/10.23919/apsipaasc44331.2018>
- Fedurek, P., Zuberbühler, K., & Semple, S. (2017). Trade-offs in the production of animal vocal sequences: Insights from the structure of wild chimpanzee pant hoots. *Frontiers in Zoology*, 14(1), 1–10. <https://doi.org/10.1186/s12983-017-0235-8>
- Fitch, W. T. (2005). The evolution of language: A comparative review. *Biology and Philosophy*, 20(2–3), 193–203. <https://doi.org/10.1007/s10539-005-5597-1>
- Fitch, W. T., De Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2(12), Article e1600723. <https://doi.org/10.1126/sciadv.1600723>
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179–210. <https://doi.org/10.1016/j.cognition.2005.02.005>
- Fitch, W. T. (2018). The biology and evolution of speech: a comparative analysis. *Annual Review of Linguistics*, 4, 255–279.

- Fridlund, A. J., & Russell, J. A. (2006). The functions of facial expressions: What's in a face? In V. Manousov, & M. Patterson (Eds.), *The Sage handbook of nonverbal communication* (pp. 299–319). Sage. <https://doi.org/10.4135/9781412976152.n16>.
- Gifford, I. I. I., Hauser, M. D., & Cohen, Y. E. (2003). Discrimination of functionally referential calls by laboratory-housed rhesus macaques: Implications for neuroethological studies. *Brain, Behavior and Evolution*, 61(4), 213–224. <https://doi.org/10.1159/000070704>
- Graham, K. E., Wilke, C., Lahiff, N. J., & Slocombe, K. E. (2020). Scratching beneath the surface: Intentionality in great ape signal production. *Philosophical Transactions of the Royal Society B*, 375(1789), 20180403. <https://doi.org/10.1098/rstb.2018.0403>
- Griebel, U., Pepperberg, I. M., & Oller, D. K. (2016). Developmental plasticity and language: A comparative perspective. *Topics in Cognitive Science*, 8(2), 435–445. <https://doi.org/10.1111/tops.12200>
- Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, 113(19), E2750–E2758. <https://doi.org/10.1073/pnas.1522072113>
- Heesen, R., Hobaiter, C., Ferrer-i-Cancho, R., & Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B*, 286(1896), 20182900. <https://doi.org/10.1098/rspb.2018.2900>
- Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. *Scientific American*, 203(3), 88–97. <https://doi.org/10.2307/24940617>. JSTOR.
- Hopkins, W. D., & Savage-Rumbaugh, E. S. (1991). Vocal communication as a function of differential rearing experiences in *Pan paniscus*: A preliminary report. *International Journal of Primatology*, 12(6), 559–583. <https://doi.org/10.1007/BF02547670>
- Ioannou, S., Chotard, H., & Davila-Ross, M. (2015). No strings attached: Physiological monitoring of rhesus monkeys (*Macaca mulatta*) with thermal imaging. *Frontiers in Behavioral Neuroscience*, 9, 160. <https://doi.org/10.3389/fnbeh.2015.00160>
- Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biology*, 10(1), 1–11. <https://doi.org/10.1186/1741-7007-10-97>
- Jhang, Y., & Oller, D. K. (2017). Emergence of functional flexibility in infant vocalizations of the first 3 months. *Frontiers in Psychology*, 8, 300. <https://doi.org/10.3389/fpsyg.2017.00300>
- Julle-Daniere, E., Micheletta, J., Whitehouse, J., Joly, M., Gass, C., Burrows, A. M., & Waller, B. M. (2015). MacFACS (Macaque Facial Action Coding System) can be used to document facial movements in Barbary macaques (*Macaca sylvanus*). *PeerJ*, 3, e1248. <https://doi.org/10.7717/peerj.1248>
- Kirchhof, J., & Hammerschmidt, K. (2006). Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*)—evidence from playback experiments. *Ethology*, 112(4), 346–354. <https://doi.org/10.1111/j.1439-0310.2006.01165.x>
- Koda, H. (2004). Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour*, 141(10), 1279–1296. <https://doi.org/10.1163/1568539042729685>
- Lameira, A. R., & Call, J. (2018). Time-space—displaced responses in the orangutan vocal system. *Science Advances*, 4(11). <https://doi.org/10.1126/sciadv.aau3401>. eaau3401.
- Lameira, A. R., Maddieson, I., & Zuberbühler, K. (2014). Primate feedstock for the evolution of consonants. *Trends in Cognitive Sciences*, 18(2), 60–62. <https://doi.org/10.1016/j.tics.2013.10.013>
- Lameira, A. R., & Shumaker, R. W. (2019). Orangutans show active voicing through a membranophone. *Scientific Reports*, 9(1), 1–6. <https://doi.org/10.1038/s41598-019-48760-7>
- Lemasson, A., Ouattara, K., Petit, E. J., & Zuberbühler, K. (2011). Social learning of vocal structure in a nonhuman primate? *BMC Evolutionary Biology*, 11(1), 362. <https://doi.org/10.1186/1471-2148-11-362>
- Locke, J. L. (2009). Evolutionary developmental linguistics: Naturalization of the faculty of language. *Language Sciences*, 31(1), 33–59. <https://doi.org/10.1016/j.langsci.2007.09.008>
- Macedonia, J. M., & Evans, C. S. (1993). Essay on contemporary issues in ethology: Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology*, 93(3), 177–197. <https://doi.org/10.1111/j.1439-0310.1993.tb00988.x>
- Maciej, P., Ndao, I., Hammerschmidt, K., & Fischer, J. (2013). Vocal communication in a complex multi-level society: Constrained acoustic structure and flexible call usage in Guinea baboons. *Frontiers in Zoology*, 10(1), 58. <https://doi.org/10.1186/1742-9994-10-58>
- Marks, M., Qiuhuan, J., Sturman, O., von Ziegler, L., Kollmorgen, S., von der Behrens, W., Mante, V., Bohacek, J., & Yanik, M. F. (2020). SIPEC: The deep-learning Swiss knife for behavioral data analysis. *bioRxiv*. <https://doi.org/10.1101/2020.10.26.355115>, 2020.10.26.355115.
- Maynard Smith, J., & Harper, D. G. C. (1995). Animal signals: Models and terminology. *Journal of Theoretical Biology*, 177(3), 305–311. <https://doi.org/10.1006/jtbi.1995.0248>
- Maynard Smith, J., & Harper, D. (2003). *Animal signals*. Oxford University Press.
- Mielke, A., Waller, B., Perez, C. J., Dubosca, J., & Micheletta, J. (2020, November 13). NetFACS: Using network science to understand facial communication systems. <https://doi.org/10.31234/osf.io/4vghk>
- Montes-Medina, A. C., Salinas-Melgoza, A., & Renton, K. (2016). Contextual flexibility in the vocal repertoire of an Amazon parrot. *Frontiers in Zoology*, 13(1), 1–13. <https://doi.org/10.1186/s12983-016-0169-6>
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist*, 111(981), 855–869. <https://doi.org/10.1086/283219>
- Müller, G. B. (2007). Evo—devo: Extending the evolutionary synthesis. *Nature Reviews Genetics*, 8(12), 943–949. <https://doi.org/10.1038/nrg2219>
- Nieder, A., & Mooney, R. (2020). The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philosophical Transactions of the Royal Society B*, 375(1789), 20190054. <https://doi.org/10.1098/rstb.2019.0054>
- Nielsen, M., & Haun, D. (2016). Why developmental psychology is incomplete without comparative and cross-cultural perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686), 20150071. <https://doi.org/10.1098/rstb.2015.0071>
- Oller, D. K. (2000). *The emergence of the speech capacity*. Routledge. <https://doi.org/10.4324/9781410602565>
- Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, 110(16), 6318–6323. <https://doi.org/10.1073/pnas.1300337110>
- Oller, D. K., Caskey, M., Yoo, H., Bene, E. R., Jhang, Y., Lee, C. C., Bowman, D. D., Long, H. L., Buder, E. H., & Vohr, B. (2019). Preterm and full term infant vocalization and the origin of language. *Scientific Reports*, 9(1), 1–10. <https://doi.org/10.1038/s41598-019-51352-0>
- Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, 32(4), 223–245. [https://doi.org/10.1016/S0021-9924\(99\)00013-1](https://doi.org/10.1016/S0021-9924(99)00013-1)
- Oller, D. K., & Griebel, U. (2014). On quantitative comparative research in communication and language evolution. *Biological Theory*, 9(3), 296–308. <https://doi.org/10.1007/s13752-014-0186-7>
- Oller, D. K., Griebel, U., Iyer, S. N., Jhang, Y., Warlaumont, A. S., Dale, R., & Call, J. (2019). Language origins viewed in spontaneous and interactive vocal rates of human and bonobo infants. *Frontiers in Psychology*, 10, 729. <https://doi.org/10.3389/fpsyg.2019.00729>
- Oller, D. K., Griebel, U., & Warlaumont, A. S. (2016). Vocal development as a guide to modeling the evolution of language. *Topics in Cognitive Science*, 8(2), 382–392. <https://doi.org/10.1111/tops.12198>
- Palagi, E., Burghardt, G. M., Smuts, B., Cordoni, G., Dall'Olio, S., Fouts, H. N., Rehakova-Petru, M., Sivi, S. M., & Pellis, S. M. (2016). Rough-and-tumble play as a window on animal communication. *Biological Reviews*, 91(2), 311–327. <https://doi.org/10.1111/brv.12172>
- Parr, L. A., Waller, B. M., Vick, S. J., & Bard, K. A. (2007). Classifying chimpanzee facial expressions using muscle action. *Emotion*, 7(1), 172. <https://psycnet.apa.org/doi/10.1037/1528-3542.7.1.172>
- Pollick, A. S., & de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184–8189. <https://doi.org/10.1073/pnas.0702624104>
- Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., & Fischer, J. (2015). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Scientific Reports*, 5(1), 1–11. <https://doi.org/10.1038/srep13220>
- Range, F., & Fischer, J. (2004). Vocal repertoire of Sooty Mangabeys (*Cercocebus torquatus atys*) in the Tai National Park. *Ethology*, 110(4), 301–321. <https://doi.org/10.1111/j.1439-0310.2004.00973.x>
- Rendall, D., & Owren, M. J. (2002). Animal vocal communication: Say what? In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 307–313). MIT Press. <https://doi.org/10.7551/mitpress/1885.001.0001>
- Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, 57(3), 583–592. <https://doi.org/10.1006/anbe.1998.1031>
- 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. *Frontiers in Psychology*, 2888.
- Robbins, M. M., Ando, C., Fawcett, K. A., Grueter, C. C., Hedwig, D., Iwata, Y., Lodwick, J. L., Masi, S., Salmi, R., Stoinski, T. S., Todd, A., Vercellio, V., & Yamagiwa, J. (2016). Behavioral variation in gorillas: Evidence of potential cultural traits. *Plos One*, 11(9), Article e0160483. <https://doi.org/10.1371/journal.pone.0160483>
- Rogers, L. J., & Kaplan, G. T. (2002). *Songs, roars, and rituals: Communication in birds, mammals, and other animals*. Harvard University Press. <https://doi.org/10.1086/394097>
- Scarantino, A., & Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour*, 100, e1–e8. <https://doi.org/10.1016/j.anbehav.2014.08.017>
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, 122, 109–116. <https://doi.org/10.1016/j.anbehav.2016.10.003>
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *Plos One*, 8(10), Article e76674. <https://doi.org/10.1371/journal.pone.0076674>
- Seyfarth, R. M., & Cheney, D. L. (2015). Social cognition. *Animal Behaviour*, 103, 191–202. <https://doi.org/10.1016/j.anbehav.2015.01.030>
- Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, 80(1), 3–8. <https://doi.org/10.1016/j.anbehav.2010.04.012>
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210(4471), 801–803. <https://doi.org/10.1126/science.7433999>

- Sievers, C., & Gruber, T. (2016). Reference in human and non-human primate communication: What does it take to refer? *Animal Cognition*, 19(4), 759–768. <https://doi.org/10.1007/s10071-016-0974-5>
- Slocombe, K. E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, 15(19), 1779–1784. <https://doi.org/10.1016/j.cub.2005.08.068>
- Slocombe, K., & Zuberbühler, K. (2010). 16. Vocal communication in chimpanzees. In E. V. Lonsdorf, S. R. Ross, & T. Matsuzawa (Eds.), *The mind of the chimpanzee* (pp. 192–207). University of Chicago Press. <https://www.degruyter.com/document/doi/10.7208/9780226492810-018/html>
- Snowdon, C., & Elowson, A. M. (2001). 'Babbling' in Pygmy Marmosets: Development after infancy. *Behaviour*, 138(10), 1235–1248. <https://doi.org/10.1163/15685390152822193>
- Stoeger, A. S., & Baotic, A. (2021). Operant control and call usage learning in African elephants. *Philosophical Transactions of the Royal Society B*, 376(1836), 20200254. <https://doi.org/10.1098/rstb.2020.0254>
- Suzuki, T. N., & Zuberbühler, K. (2019). Animal syntax. *Current Biology*, 29(14), R669–R671. <https://doi.org/10.1016/j.cub.2019.05.045>
- Taylor, D., Dezecache, G., & Davila-Ross, M. (2021). Filling in the gaps: Acoustic gradation increases in the vocal ontogeny of chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, 83(5), Article e23249. <https://doi.org/10.1002/ajp.23249>
- Taylor, A. M., Reby, D., & McComb, K. (2009). Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethology*, 115(10), 905–915. <https://doi.org/10.1111/j.1439-0310.2009.01681>
- Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., Goncalves, I. B., Burkhart, J. M., Flower, T., Gaunet, F., Glock, H. J., Gruber, T., Jansen, D. A. W. A. M., Liebal, K., Linke, A., Miklósi, A., Moore, R., van Schaik, C. P., Stoll, S., ... Manser, M. B. (2017). Exorcising Grice's ghost: An empirical approach to studying intentional communication in animals. *Biological Reviews*, 92(3), 1427–1433. <https://doi.org/10.1111/bvr.12289>
- Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: The past, present and the future. *Ethology*, 119(1), 1–11. <https://doi.org/10.1111/eth.12015>
- Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015). Characterizing vocal repertoires—Hard vs. soft classification approaches. *PLoS One*, 10(4), e0125785.
- Wathan, J., Burrows, A. M., Waller, B. M., & McComb, K. (2015). EquiFACS: The equine facial action coding system. *PLoS One*, 10(8), Article e0131738. <https://doi.org/10.1371/journal.pone.0131738>
- Watts, D. P. (2016). Production of grooming-associated sounds by chimpanzees (*Pan troglodytes*) at Ngogo: Variation, social learning, and possible functions. *Primates*, 57(1), 61–72. <https://doi.org/10.1007/s10329-015-0497-8>
- Weary, D. M., Jasper, J., & Hötzel, M. J. (2008). Understanding weaning distress. *Applied Animal Behaviour Science*, 110(1), 24–41. <https://doi.org/10.1016/j.applanim.2007.03.025>
- Weierich, M. R., Wright, C. I., Negreira, A., Dickerson, B. C., & Barrett, L. F. (2010). Novelty as a dimension in the affective brain. *NeuroImage*, 49(3), 2871–2878. <https://doi.org/10.1016/j.neuroimage.2009.09.047>
- Wich, S. A., Krützen, M., Lameira, A. R., Nater, A., Arora, N., Bastian, M. L., Meulman, E., Morrogh-Bernard, H. C., Atmoko, S. S. U., Pamungkas, J., Perwitasari-Farajallah, D., Hardus, M. E., van Noordwijk, M., & van Schaik, C. P. (2012). Call cultures in orang-utans? *PLoS One*, 7(5), Article e36180. <https://doi.org/10.1371/journal.pone.0036180>
- Zuberbühler, K. (2011). Cooperative breeding and the evolution of vocal flexibility. In K. R. Gibson, & M. Tallerman (Eds.), *The Oxford handbook of language evolution* (pp. 71–83). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199541119.013.0005>
- Zuberbühler, K. (2018). Combinatorial capacities in primates. *Current Opinion in Behavioral Sciences*, 21, 161–169. <https://doi.org/10.1016/j.cobeha.2018.03.015>
- Zuberbühler, K. (2020). Syntax and compositionality in animal communication. *Philosophical Transactions of the Royal Society B*, 375(1789), 20190062. <https://doi.org/10.1098/rstb.2019.0062>