

Call type signals caller goal: a new take on ultimate and proximate influences in vocal production

Isaac Schamberg^{1*} , Roman M. Wittig^{1,2} and Catharine Crockford^{1,2} 

¹*Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*

²*Tai Chimpanzee Project, Centre Suisse de Recherches Scientifiques, Abidjan 01, Côte D'ivoire*

ABSTRACT

After 40 years of debate it remains unclear whether signallers produce vocalizations in order to provide receivers with information about call context or external stimuli. This has led some researchers to propose that call production is arousal- or affect-based. Although arousal influences certain acoustic parameters *within* a call type, we argue that it cannot explain why individuals across vertebrates produce *different* call types. Given emerging evidence that calls are goal-based, we argue that call type is a signal of a caller's goal to elicit a change in receiver behaviour. Using chimpanzees (*Pan troglodytes*) and vervet monkeys (*Cercopithecus aethiops*) as case studies, we demonstrate the two benefits of viewing call production as signalling both caller goal (which determines call type) and caller arousal (which affects within-call-type variation). Such a framework can explain first, why a single class of calls is apparently given in multiple contexts, and, second, why some species have larger call repertoires than others. Previous studies have noted links between sociality and repertoire size, but have not specified exactly why animals living in societies that are more complex might require a greater number of differentiated signals. The caller-goal framework potentially clarifies how social complexity might favour call diversification. As social complexity increases, callers may need to elicit a larger number of distinct behaviours from a wider range of distinct audiences.

Key words: vocal production, call diversity, animal communication, communication goals, chimpanzee.

CONTENTS

I. Introduction	2
II. Caller arousal	3
III. Caller goal	5
IV. Signal differentiation	6
V. Chimpanzee call production	7
VI. Vervet call production	8
VII. Discussion	9
(1) Caller goal and call function	9
(2) Signal differentiation	9
(3) Cognition and communication	9
VIII. Conclusions	10
IX. Acknowledgments	10
X. References	10

* Address for correspondence (Tel: +49 341 3550 220; E-mail: isaac.schamberg@gmail.com).

I. INTRODUCTION

What causes an individual animal to produce a particular call type at a particular moment? This question remains unanswered after more than 40 years of investigation into mammalian communication. Much research has documented the ‘referential’, or context-specific, nature of vocalizations – i.e. certain call types are produced in a relatively narrow set of circumstances and receivers are able to extract specific information about call context (Macedonia & Evans, 1993; Manser, 2010; Wheeler & Fischer, 2012; Townsend & Manser, 2013). Although call context often correlates with call type, it is far from a one-to-one correspondence. Even highly context-specific calls are sometimes produced outside their primary context [e.g. vervet monkey (*Cercopithecus aethiops*) ‘snake alarm calls’ are also produced during intergroup encounters (Strusaker, 1967; Price *et al.*, 2015)]. Additionally, callers sometimes produce different call types in the same context [e.g. chimpanzees (*Pan troglodytes*) produce both hoos and barks upon detection of a predator (Schel *et al.*, 2013b)]. Such messy patterns of call production suggest that calls may not ‘refer’ to external stimuli or objects (Snowdon, 1990; Wheeler & Fischer, 2012). Rather, call production appears to be the result of a process intrinsic to the caller (Fischer & Price, 2016). Researchers have variously characterized this enigmatic internal process as ‘arousal’, ‘emotion’, ‘affect’, ‘motivational state’, or simply ‘internal state’ (Owren & Rendall, 2001; Rendall, Owren, & Ryan, 2009; Wheeler & Fischer, 2012; Fischer & Price, 2016). Although these terms refer to distinct phenomena, all highlight the involuntary nature of call production, and (at least implicitly) draw attention to the discrepancy between signallers and receivers: vocalizations merely burst forth from signallers, whereas receivers exhibit considerable sophistication and flexibility when interpreting these calls by integrating various sources of information (Seyfarth & Cheney, 2018; Wheeler & Fischer, 2012; Fischer, 2017; but see Bar-On & Moore, 2017; Cheney & Seyfarth, 2018).

This apparent chasm between signallers and receivers, however, may be an artifact of a recent focus on the context-specific nature of call production rather than a reflection of the true mechanisms underlying the two processes (Bar-On & Moore, 2017; Gruber & Grandjean, 2017; Cheney & Seyfarth, 2018). When viewed through the lens of context-specificity, call production appears unsophisticated because callers may not intend to provide the information that receivers extract when hearing calls. For example, receivers hear an alarm call and infer there is an eagle overhead, but evidence suggests that callers may produce such calls in response to an internal state brought about by the presence of the eagle, rather than producing the calls to convey information about the eagle *per se* (Seyfarth, Cheney, & Marler, 1980; Price *et al.*, 2015). A context-specific framework tackles the issue of call production by considering a signal’s benefit to a receiver. From the signaller’s perspective, however, providing information

about call context may not be the primary benefit of vocalizing.

At an ultimate evolutionary level, signallers vocalize in order to influence the behaviour of receivers, and signals become stable within a repertoire when the signal confers a mutual benefit to both signaller and receiver (Krebs & Dawkins, 1984). For example, threat displays, common across animals, convey to receivers imminent escalation of aggression. Receivers then have a choice to respond either by retreating or engaging, likely limiting the number of costly fights in which both signaller and receiver engage (Laidre & Johnstone, 2013). How this function is carried out *via* proximate mechanisms remains a central question for scholars of vocal communication. One commonly proposed explanation is that calls are affect-based – i.e. calls reflect the internal emotional state of the caller (Owren & Rendall, 2001; Rendall *et al.*, 2009). Under such a framework, the ultimate function of a scream might be to deter further aggression, but, at a proximate level, screams are the result of acute fear. Another possibility is that the proximate mechanism of call production reflects the ultimate function. For instance, at an ultimate level, a long-call may function to attract social partners to a caller, and an individual’s proximate motivation for producing such a call may also be a desire to attract particular social partners, reflecting the signal’s evolutionary function.

Here, we consider the function of a call to be an individual’s communicative goal, such that signallers produce a given call type because they are motivated at that moment to achieve a particular goal, and vocalizing is one way to realize that goal. (Owings & Morton, 1998; Morton, 2017). Our use of the term ‘goal’ does not necessarily imply that a caller has any (implicit or explicit) understanding of how or why the signal influences receiver behaviour in the manner the caller desires. Thus, our conception of caller goal differs from the ‘goal-directed’ vocalizations discussed in conjunction with intentional signalling, despite the centrality of a caller goal to both. A caller, for example, might produce a goal-based vocalization without any ‘intentionality’, if the caller does not modify or elaborate the signal according to the receiver’s behaviour, a key criterion for intentionality. For us, the term ‘caller goal’ merely signifies that a caller has a motivation that a receiver behaves in a certain way. Other researchers use the term to signify a signaller’s awareness or intention of how the signal can be used to change the receiver’s behaviour (*sensu* Dickinson & Balleine, 2000; Townsend *et al.*, 2017). Unlike these discussions of intentionality, herein ‘caller goal’ is independent of any meta-awareness or flexibility in vocal production. In the threat display example, whether the display is a given by a crab or a human, the proximate goal – and, indeed, the ultimate function – is likely the same: deterrence of an intruder. The difference between the signals is the extent to which the signallers understand the effect the display has on receivers. A human may understand that the display scares the intruder by making the displayer appear larger than he really is. The crab, by contrast, may only perceive that intruders often flee when the signaller performs

the display. Despite this important difference between the two displays, we consider them both to be motivated by a signaller's goal of receiver deterrence. Hence, our broad conception of 'caller goal' encompasses both biological and psychological definitions. The biological definition is that the signaller simply behaves to achieve a fitness-enhancing goal that happens to require changing the behaviour of another animal (*sensu* Mayr, 1961). Psychological definitions include that behaviour is at varying levels imbued with the signaller's intention to change the receiver's behaviour (*sensu* Dickinson & Balleine, 2000; Townsend *et al.*, 2017).

Our two alternative explanations for call production – affect- or goal-based call production – are not mutually exclusive, but they do represent two potentially different explanations for call production. Here, our aim is to clarify the contributions of affect and caller goal on call production by demonstrating that affect can account for certain features of call production (a scream may be higher or lower pitched), but not for the production of specific call types (the emission of a scream rather than a grunt), which are best explained by caller goal. If this premise is correct, then at both the ultimate and proximate levels, each call type in a species' vocal repertoire is designed to elicit a specific behaviour from receivers (such as 'approach', 'retreat', 'mate'), or from a specific subset of receivers (such as 'females' or 'kin'). Emotional valence and urgency embedded within each call type add more-specific information for receivers (such as 'approach cautiously' or 'retreat fast').

The framework we propose echoes, in certain ways, what many theorists and researchers have already pointed out: vocalizations convey information about both caller emotion (or arousal), on the one hand, and 'referential' (or semantic) information about the environment on the other (Premack, 1972; Seyfarth & Cheney, 1997; Manser, 2001). Previous formulations of this notion, however, frequently have been framed from the perspective of receivers. Premack (1972), for example, argued that the mechanisms underlying call production might be entirely 'emotional', but receivers would be able to extract non-emotive information about context from such calls so that for receivers calls might function as if they are referential (Macedonia & Evans, 1993). Our aim here is to address this arousal–referential duality from the caller's perspective and argue that, just as receivers extract emotional and non-emotional information from calls, both emotional and non-emotional mechanisms shape call production as well.

Here, by referring to examples across taxa, we present evidence that while caller arousal affects call prosody within call types, different call types signal the caller's goal. Finally, we demonstrate how the caller goal framework might operate within a vocal repertoire by drawing on well-studied examples in both chimpanzees (*Pan troglodytes*) and vervet monkeys (*Chlorocebus aethiops*), and hence highlight the insights this framework can provide.

II. CALLER AROUSAL

In linguistics, *prosody* refers to the acoustic measurements associated with speaker emotion or arousal (Cutler, Dahan, & Van Donselaar, 1997). Emotion is well conveyed in speech without necessarily changing the words emitted. Rather, the prosodic features of the words are altered, where prosodic features are a standard set of acoustic measures that characterise the pitch, rhythm, amplitude and tonality of words (Hirst & Di Cristo, 1998). We will use the term *call prosody* to refer to the acoustic measurements related to caller arousal or affective state. We believe that importing this term from linguistics into animal communication is warranted because, across taxa, certain acoustic features correlate with caller arousal just as they do in humans (Table 1; see Briefer, 2012, for review). While differences between the acoustic features that correlate with caller arousal exist among species, there is a clear across-taxa pattern: high-arousal callers produce noisier calls with higher frequencies, at faster rates or with longer duration than do low-arousal callers. Red-fronted lemurs (*Eulemur rufifrons*), for example, produce the same call type ('woof') in two contexts that elicit differential arousal levels: during intergroup encounters (high arousal) and upon sighting domestic dogs (low arousal). Woofs produced during intergroup encounters had more energy in higher frequency bands – i.e. had a higher frequency – than woofs produced in response to dogs (Fichtel & Hammerschmidt, 2002). In meerkat (*Suricata suricatta*) alarm calls, caller arousal is expressed *via* a vocalization's harmonic-to-noise ratio – i.e. the noisiness of a call. Callers closer to predators produce noisier calls compared to more distant callers (Manser, 2001).

In addition to the parameters of frequency and noisiness, call rate, inter-call interval, and call duration also commonly correlate with caller arousal (reviewed in Briefer, 2012). For example in chacma baboons (*Papio ursinus*), Rendall (2003) showed that within a context, specifically 'move' or 'social approach' contexts, baboons in a high-arousal rather than a low-arousal state, produced more grunts with shorter inter-call intervals that were higher pitched and noisier. Spotted hyenas (*Crocuta crocuta*), too, produce the same call type ('whoops') in low-arousal situations (when resting) as well high-arousal situations (during conflicts with lions) (Gersick *et al.*, 2015). Australian fur seals (*Arctocephalus pusillus doriferus*) in high- rather than low-arousal contexts produce higher frequency barks at faster rates (Tripovich *et al.*, 2008). In one experiment, chimpanzees were exposed to two experimental conditions: one in which they heard the pant hoot of a familiar individual, another in which they heard the pant hoot of a stranger. These two conditions likely elicited divergent levels of arousal in the subjects given that chimpanzees are extremely hostile towards out-group members. Nevertheless, subjects produced the same call type despite substantial differences in arousal (Kajikawa & Hasegawa, 2000). This pattern is consistent for many of the examples in Table 1. The effect that caller arousal has on call production can be seen in variation in call frequency, rhythm, and noisiness within a particular call type. However,

Table 1. Acoustic features associated with caller arousal. Arrows indicate whether the acoustic property increases or decreases with caller arousal. F0: fundamental frequency; Amp: amplitude; Dur: duration.

Species	Authors	Call type	F0	Noisiness	Dur	Amp	Call rate
Primates							
Marmoset (<i>Callithrix j. jacchus</i>)	Schrader & Todt (1993)	Contact calls	↑				
Squirrel monkey (<i>Saimiri sciureus</i>)	Fichtel, Hammerschmidt, & Jürgens (2001)	Various	↑				
Red-fronted lemur (<i>Eulemur rufifrons</i>)	Fichtel & Hammerschmidt (2002)	Woofs	↑				
Barbary macaque (<i>Macaca sylvanus</i>)	Fischer, Hammerschmidt, & Todt (1995)	Disturbance calls	↑				
Pigtail macaque (<i>Macaca nemestrina</i>)	Gouzoules & Gouzoules (1989)	Screams	↑				
Chacma baboon (<i>Papio ursinus</i>)	Rendall (2003)	Grunts	↑	↑	↑		↑
Chacma baboon (<i>Papio ursinus</i>)	Fischer <i>et al.</i> (2001)	Barks		↑			
Bonnet macaque (<i>Macaca radiata</i>)	Cross, McCowan, & Ramakrishnan (2007)	Alarm calls		↑			
Human (<i>Homo sapiens</i>)	Ruiz <i>et al.</i> (1996)	Talking	↑				
Chimpanzee (<i>Pan troglodytes</i>)	Slocombe & Zuberbühler (2007)	Screams	↑		↑		
Chimpanzee (<i>Pan troglodytes</i>)	Kajikawa & Hasegawa (2000)	Pant hoots	↑		↓		↓
Other mammals							
Tree shrew (<i>Tupaia belangeri</i>)	Schekka, Esser, & Zimmermann (2007)	Squeaks	↑				↑
Bottlenose dolphin (<i>Tursiops truncatus</i>)	Esch <i>et al.</i> (2009)	Whistles	↑				↑
Cape ground squirrel (<i>Xerus inauris</i>)	Furrer & Manser (2009)	Alarm calls	↑		↑		
African elephant (<i>Loxodonta Africana</i>)	Soltis <i>et al.</i> (2009)	Rumbles	↑			↑	
Australian fur seal (<i>Arctocephalus pusillus doriferus</i>)	Tripovich <i>et al.</i> (2008)	Barks	↑				↑
Dwarf mongoose (<i>Helogale parvula</i>)	Rubow, Cherry, & Sharpe (2017)	Recruitment calls					↑
Spotted hyena (<i>Crocuta crocuta</i>)	Gersick <i>et al.</i> (2015)	Whoops					↑
Meerkat (<i>Suricata suricatta</i>)	Manser (2001)	Alarm calls		↑			

in both high- and low-arousal contexts, callers often produced the same call type (e.g. in both arousal states, baboons produce grunts, seals produce barks, and hyenas produce whoops).

Specifying the acoustic parameters associated with caller arousal that vary *within* a call type and using the term *call prosody* to refer to them may be useful in clarifying the influence of caller arousal when emitting a particular call type. Note that some of the acoustic features that are associated with call prosody are also used to discriminate different call types, for example, differences in context-specific bark variants in chimpanzees can be

discriminated using call duration and pitch (Crockford & Boesch, 2003). However, additional acoustic features, such as those describing the frequency or spectral distribution, are also valuable for discriminating between call types [e.g. baboon bark variants (Fischer *et al.*, 2001); baboon grunt variants (Rendall *et al.*, 1999)].

In summary, when arousal or affect is discussed in relation to call production, readers are sometimes left with the impression that caller arousal is the primary influence of call production (Owren, Amoss, & Rendall, 2011; Wheeler & Fischer, 2012). The evidence suggests that caller arousal affects certain parameters of call structure that shape

variation *within* a call type, but may not explain the motivation that prompts an individual to produce different call types.

III. CALLER GOAL

As argued above, caller arousal alone is unlikely to explain the production of particular call types. If caller arousal does not explain production of a particular call type, what does? We propose that a caller's goal is the primary determinant of call type: call type signals caller goal, and different call types signal different goals (using our broad definition of 'goal' as stated above; Fig. 1). While this theory has not been directly tested, there is mounting evidence that callers' goals are an important feature of call production (Genty *et al.*, 2014; Fischer & Price, 2016; Gruber & Grandjean, 2017; Townsend *et al.*, 2017).

Earlier investigations into primate vocalizations did not typically address whether calls are goal-based, but the existence of audience effects – the phenomenon of audience composition or audience behaviour affecting signalling behaviour – offered some support to the notion of goal-based vocalizations (reviewed in Fichtel & Manser, 2010). Signallers seem to alter their call production depending on attributes of potential receivers, such as their relationship to the caller, distance from the audience to a threat, or the potential information available to the audience about a threat (Cheney & Seyfarth, 1985; Papworth *et al.*, 2008; Crockford *et al.*, 2012; Crockford, Wittig, & Zuberbühler, 2017). Upon seeing a silhouette of a predator, male chickens (*Gallus gallus*) are more likely to emit context-specific alarm calls if there is a potential mate present rather than another male chicken or a female of another species (Karakashain, Gyger, & Marler, 1988). Pygmy marmosets (*Cebuella pygmaea*) alter the acoustic structure of their contact calls based on the distance of their audience (Snowdon & de la Torre, 2002). Similarly, vervet mothers are more likely to produce alarm calls when with their own offspring than with unrelated juveniles (Cheney & Seyfarth, 1985). Chimpanzees at a food patch emit context-specific food-associated grunts that function to recruit conspecifics to food (Kalan & Boesch, 2015). They are more likely to emit food-associated grunts when a simulated friend or dominant individual is nearby compared to other simulated group members (Schel *et al.*, 2013a). Chimpanzees also produce more long-distance pant hoots when closely bonded individuals are out of sight, but within earshot, than when closely bonded individuals are either present and visible, or are out of audible range (Mitani & Nishida, 1993).

Furthermore, callers can also withhold vocalizations if being overheard might precipitate unwarranted receiver behaviour. Male chickens, for example, are less likely to give vocal displays to females when rival males are present than when not (Smith, Taylor, & Evans, 2011). Female chimpanzees appear to suppress copulation calls when copulating with low-ranking males when high-ranking

males are nearby (Townsend, Deschner, & Zuberbühler, 2008; Fallon *et al.*, 2016), presumably to facilitate female choice.

Such examples are consistent with the notion that callers produce vocalizations in order to elicit a certain behaviour from a particular audience, rather than vocalizing solely out of excitement, fear, or arousal. If calls are produced with the goal of influencing the behaviour of specific receivers, callers should be sensitive to who will hear a vocalization, which is precisely what studies on audience effects find. It is possible, though, to offer explanations of audience effects *via* the effect an audience has on a signaller's arousal (Owren *et al.*, 2011). For example, proximity to one's own offspring in the presence of a predator might cause heightened arousal, and therefore result in increased call production. However, one could easily make the opposite prediction: vervet mothers would be more agitated if they could not see their offspring during a dangerous predator sighting, and would thus be more likely to call in the absence of their offspring. Similarly for the chimpanzee example, one could argue that the absence of a closely bonded individual causes increased arousal in potential signallers, but one could equally easily argue that the presence of an ally would increase arousal. An arousal-based argument for the audience effect reported by Mitani & Nishida (1993) is even more problematic because the call production is not based on mere presence or absence, but on whether the bonded individual is out of sight, but within earshot.

Recently, several studies have addressed the question of caller goal more directly by demonstrating persistence in call production – that is, call production that continues until a receivers' behaviour has changed, apparently satisfying a caller's putative goal. For example, in Thomas langurs (*Presbytis thomasi*), the dominant male produces alarm calls until all nearby individuals have also produced at least one alarm call (Wich & de Vries, 2006). Such a pattern of call production possibly suggests that the dominant male produces alarm calls with the goal of eliciting alarm calls from other group members. Similarly, bonobos (*Pan paniscus*) produce two call combinations that exhibit call persistence. In one case, individuals continue to call until they receive a response from an out-of-sight individual, suggesting that the goal of the signal is to elicit a vocal response from an out-of-sight group mate (Schamberg *et al.*, 2016). In the other case, bonobos continue to call until individuals from outside their own party approach them, supporting the hypothesis that individuals are calling with the aim of recruiting individuals to their own party (Schamberg *et al.*, 2017). Finally, chimpanzees produce certain alarm calls persistently, ceasing only when audience members are out of danger, suggesting the goal of the caller is to warn group members (Schel *et al.*, 2013b; Crockford *et al.*, 2017). A framework for testing whether a call is intentionally emitted to change another's behaviour ('first-order' intentionality; Dennett, 1983) is provided by Townsend *et al.* (2017); see Genty *et al.* (2014) and Gruber & Grandjean (2017) for helpful discussions on this topic.

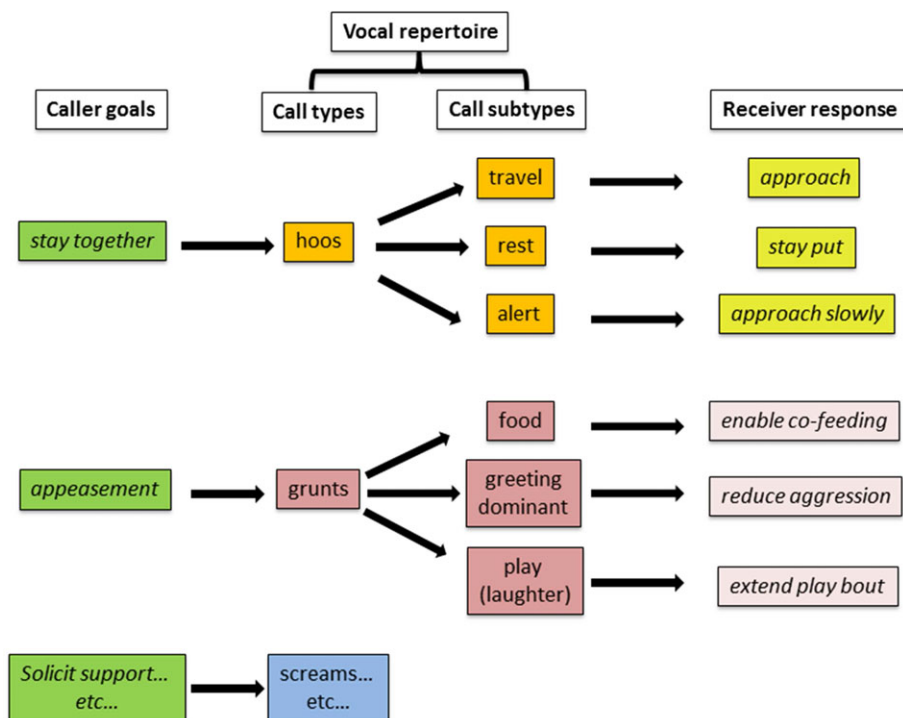


Fig. 1. Schematic of part of the chimpanzee vocal repertoire using the caller-goal framework. Different caller goals promote usage of different call types. Call types may have acoustically distinguishable subtypes each requiring a different response from receivers to achieve the caller goal. Here, call subtypes are labelled according to the contexts in which they are produced. Contextual information likely enables receivers to respond appropriately, and we predict should be more likely to be conveyed in the acoustic structure of the calls when visual cues are restricted, such as in low-visibility habitat.

Vocal-production patterns of signallers across a broad range of taxa – including chickens and primates – indicate that a particular call type is not only emitted in a specific context – such as predator, food, mate attraction or contact maintenance – but depends on the audience that can hear it. In some cases, callers apparently persist in calling until a certain change in receiver behaviour has been elicited. Putting these various threads of evidence together, it would seem that particular call types function to elicit a certain behaviour that fulfils a caller's goal. The fact that vocal production is influenced by audience effects suggests that signallers direct calls to receivers from whom eliciting a particular response would be beneficial, and withhold calls when a particular response from receivers could be costly.

Whilst some calls are only produced in specific contexts, such as food-associated calls (Schel *et al.*, 2013a), or specific alarm calls produced in response to particular predators (Cheney & Seyfarth, 1990), other calls can be relatively non-specific with regards to context, with apparently the same call being produced in several contexts, such as species where calls produced in response to predators are also emitted in aggression contexts (e.g. Wheeler & Fischer, 2012). Receivers may nonetheless infer situational information from the call due to associated contextual cues (e.g. Wheeler & Fischer, 2012; Cheney & Seyfarth, 2018): receivers can see whether there is a predator or a conspecific aggressor (e.g. Wheeler & Fischer, 2012; Seyfarth & Cheney, 2018). Here,

we posit that when contextual cues are unreliable, such as for species living in low-visibility habitat, then in order for signallers to elicit desired receiver responses reliably, selection may act on call structure to encode contextual information.

The caller-goal framework draws on biological principles linking call function to caller goal. In addition to call function, caller goal takes audience effects into account, such that callers target vocalizations at particular audiences (mates, kin) presumably with the goal (intentional or unintentional) of evoking a particular response from receivers.

IV. SIGNAL DIFFERENTIATION

Why do some species have larger call repertoires than others? Reframing the study of call production around caller goals may prove useful in answering this question. If it benefits signallers to elicit a greater range of responses from receivers, would we expect these animals to have a larger vocal repertoire than others? The link between selection pressure and call repertoire size is clearest – and perhaps most discussed – in the case of predation pressure. Species vulnerable to a variety of predators that require distinct escape strategies have generally evolved call types that correspond to these different predators (Furrer & Manser, 2009). The fitness consequences for receivers of engaging in

the wrong escape response are significant. It is, therefore, unsurprising that natural selection would favour signals that elicit specific responses in these contexts (e.g. Seyfarth *et al.*, 1980; Price *et al.*, 2015). Predation, however, is not the only fitness-impacting context that has the potential to shape call repertoires. The formation and maintenance of social bonds has numerous links to fitness (e.g. offspring survival, longevity, siring of offspring, ability to respond to stressful events, rank acquisition) (Silk, Alberts, & Altmann, 2003; Schülke *et al.*, 2010; Silk *et al.*, 2010; Silk, Seyfarth, & Cheney, 2016; Wittig *et al.*, 2016). Natural selection, then, might be expected to favour specific signals for coordinating distinct behaviours in the social domain as well.

Although it is difficult accurately to compare call-repertoire size across species, several studies have found a correlation between various measures of sociality and call-repertoire size, but the reason for such a link is not clear (Blumstein & Armitage, 1997; McComb & Semple, 2005; Blumstein, 2007; Gustison, le Roux, & Bergman, 2012). In an excellent review on the relationship between social and communicative complexity, Pollard & Blumstein (2012) propose several features of sociality that may impose selection pressures on individuals' signalling behaviour – namely, that signallers living in large groups with highly differentiated relationships and numerous distinct demographic roles may have 'more to talk about' than signallers living in smaller, less-complex groups. While intuitively appealing, the notion of 'more to talk about' is vague. A goal-based theory of call production offers a clearer formulation of the same idea. Specifically, individuals living in complex societies may have 'more to talk about' because they have a larger number of distinct responses to elicit from receivers. Callers that interact with more individuals and have more differentiated relationships with these individuals (e.g. bonded *versus* non-bonded; sexual partner *versus* sexual rival; kin *versus* non-kin; dominant *versus* subordinate; dependent *versus* independent) require different responses from these different receivers. Additionally, callers that inhabit a number of distinct cooperative roles (e.g. coalition partner, monogamous pair, caretaker, food sharer, sentinel, hunters or territory defenders) may also benefit from eliciting a variety of distinct receiver responses. Eliciting specific coordinated or cooperative behaviour may require specific signals, such as is evident in the cooperatively breeding mongooses, but not in the sympatric ground squirrels exposed to the same predators (Furrer & Manser, 2009). Thus, when considering the relationship between sociality and communication, one central factor to consider may be how sociality affects the potential goals that signallers have when producing different vocalizations.

A large number of specific receiver responses may be necessary but not sufficient for the evolution of large call repertoires. There is ample evidence that receivers are quite sophisticated and flexible when responding to calls (reviewed in Seyfarth *et al.*, 2010; Fischer & Price, 2016; Cheney & Seyfarth, 2018). In many cases, then, general call types may suffice to coordinate specific behaviours. However, in situations where selection pressure is especially strong (e.g.

predator detection or cooperation) or contextual information is unreliable (e.g. low-visibility habitat), we expect to observe signal differentiation of general call types into subtypes, where each call subtype is associated with a particular receiver response but is driven by the same over-arching caller goal. Thus, larger call repertoires should be expected in social systems where callers benefit from eliciting a large number of specific receiver responses, and contextual cues for receivers are limited due to low visibility or long distances between callers and receivers. Note that larger repertoires may mean producing more distinct call types, more call subtypes or more call combinations. Constraints of anatomy as well as of speciation (enabling individuals to reliably recognize calls from their own rather than from another species), likely limit the number of distinct call types a species can emit, but this is unlikely to be the limiting factor in terms of the capacity to convey different information. Most of the 7000 extant human languages (Simons & Fennig, 2018) use no more than 50–100 unique sounds or phonemes (Crystal, 2010). This limited set of phonemes, however, is combined to make endless novel meanings.

The caller-goal framework provides a clear rationale as to when vocal differentiation within a vocal repertoire would be expected. Under conditions where callers require many rather than few different receiver responses to achieve their goals, such as for individuals living in more complex societies, selection pressures should act to produce greater vocal differentiation, with more acoustically distinguishable calls. In accordance with the social complexity hypothesis, we would expect callers living in more-complex societies (e.g. McComb & Semple, 2005), with differentiated relationships and more social roles, to require more receiver responses from more receivers to achieve their goals. If so, we would expect greater vocal differentiation in these societies, especially when living in environments where contextual information through non-vocal channels is restricted. We explore the potential use of a caller-goal framework to clarify apparently messy vocal systems and to aid understanding of call differentiation. We use the vocal repertoires of two relatively well-studied species, chimpanzee and vervet monkey, which both contain examples of apparently the same or similar calls being emitted in apparently disparate contexts.

V. CHIMPANZEE CALL PRODUCTION

Chimpanzees, like many species, have a vocal repertoire that consists of a small number of general call types (e.g. hoos, grunts, barks and screams) that can be subdivided into a larger number of call subtypes (Goodall, 1986; Crockford & Boesch, 2003, 2005; Slocombe & Zuberbühler, 2007). Call subtypes within the same class of calls are sometimes produced in very different contexts. For example, chimpanzees produce at least three types of quiet 'hoo' calls in different contexts (travel, rest, and alert) (Gruber & Zuberbühler, 2013; Crockford, Wittig, & Zuberbühler, 2015; Crockford *et al.*, 2012, 2017;

Crockford, Gruber & Zuberbühler, 2018). It is unlikely that these different contexts elicit similar arousal levels in individuals. Thus, explanations for call production based on context or arousal seem unlikely to explain why callers produce similar calls in such a wide range of situations. A caller-goal framework of call production presumes that all hooos have a common, underlying goal. For chimpanzee hooos, playback and observational studies suggest that the signaller's overarching goal is to maintain cohesion (Gruber & Zuberbühler, 2013; Crockford *et al.*, 2015; Crockford *et al.*, 2017). In order to maintain cohesion, however, receivers must respond differently to signallers in each context: in rest contexts, receivers must stay put in the vicinity of signallers; in travel contexts, receivers need to approach signallers; and in alert contexts such as seeing snakes, receivers must approach signallers slowly with caution (Crockford *et al.*, in press). Given that receivers often cannot see camouflaged threats until they are close to them, call types that differentiate whether a fast or cautious approach is required are likely adaptive to both callers and receivers (Fig. 1). Note here that whilst maintaining cohesion may be the overarching goal, more-specific goals are also evident. In travel contexts, callers are presumably not only motivated to be together, but specifically to travel together. In alert contexts, additional motivation is presumably to warn receivers of a threat.

For species living in habitats with good visibility, contextual cues are available to receivers in addition to the acoustic properties of the call (Wheeler & Fischer, 2012). It is thus possible that a single hoo type could elicit context-appropriate receiver responses (i.e. staying put, approaching fast, or approaching cautiously) if receivers are able to use contextual information to determine the best response. Both chimpanzee habitat and social structure, however, render it difficult for receivers to rely on such cues. Chimpanzees generally live in low-visibility forests where signallers and receivers are out-of-sight of one another. Visual cues that might provide additional information to that conveyed by a call type are often unavailable, even for individuals who are 20 m apart. Also, chimpanzees live in a fission–fusion society, such that individuals within a community can be separated for hours, days or weeks at a time. At any given time, some individuals may be feeding or resting, while others begin to travel to a distant location. When subgroup composition changes frequently and unpredictably as in chimpanzee society it is easy for an individual to become separated from an important ally. While separation from an ally may not immediately seem to be a fitness-affecting event, individuals that maintain bond partners likely gain direct benefits from supporting each other during within- and between-group conflict (Duffy, Wrangham, & Silk, 2007; Langergraber, Mitani, & Vigilant, 2007; Wittig *et al.*, 2016). This benefit is only available when allies' movements are coordinated. Thus, for individuals who live in fluid social groups and who accrue fitness benefits from maintaining cooperative relationships, vocalizations that facilitate coordination may be under positive selection. In addition, when contextual information

is not otherwise available, signals that elicit specific responses (e.g. stay, approach quickly, or approach cautiously) may better facilitate coordination – and cooperation – between allies, creating positive selection for signal differentiation.

A caller-goal framework may also help build hypotheses about the function of call types and assist in understanding the structure and function of vocal repertoires in general. Chimpanzees, for example, have several context-specific grunt vocalizations, including 'pant grunts', 'laughter', and 'rough grunts'. Neither context, arousal nor affect can explain why callers produce grunts in such seemingly disparate situations as reunion with a dominant individual (pant grunts), play (laughter), and feeding (rough grunts). Under a goal-based framework, however, a general call class is assumed to have a common goal, and one can hypothesize about the goal of grunts (Fig. 1). Indeed, it seems likely that the general goal might be to appease the target of the vocalization in order to reduce the risk of aggression. Reunions with dominant individuals and play are both contexts where callers may benefit from producing an appeasement signal to ensure that the interaction remains peaceful. Callers may produce rough grunts while feeding in order to signal a willingness to co-feed with other individuals and avoid aggression. As in the case of hooos, a single call type may have evolved into several call subtypes in order to elicit specific behaviours from receivers. The goal of laughter is likely to elicit further play; of pant grunts to suppress aggression from dominants; and of rough grunts to facilitate co-feeding.

VI. VERVET CALL PRODUCTION

Having examined parts of the chimpanzee vocal repertoire using the caller-goal framework, we now do the same with well-studied vervet monkey vocalizations (e.g. Seyfarth *et al.*, 1980; Cheney & Seyfarth, 1990). Vervet monkeys also employ alarm calls for a range of different predators, although their savannah–woodland habitats are more open and visibility is better. A recent reanalysis of vervet monkey alarm calls found that the calls that females produced in response to snakes and eagles ('chutters' and 'rraups', respectively) were acoustically similar to those given during bouts of within- and between-group aggression (Price *et al.*, 2015). The authors concluded that chatter and rraup calls produced primarily to snakes and eagles, respectively, cannot refer to snakes or eagles in a linguistic sense because they would not then also produce these calls in the context of within- and between-group aggression. If the calls are not referential signals, what is it that they signal? One possibility is that predator sightings and aggressive context encounters elicit similar affective states, which cause callers to produce the same calls. Such an explanation may be plausible, but it struggles to address certain features of the call production. First, predation events and aggression may pose different threats to individuals and thus seem unlikely to elicit the same affective state. Second, as mentioned earlier,

an arousal-based explanation cannot explain why vervets are more likely to call in response to a leopard in the presence of nearby kin (Cheney & Seyfarth, 1985). In a goal-based framework, chutters and rraups may signal particular goals that callers have in multiple contexts. For example, signallers may seek to elicit similar behaviour from conspecifics in eagle and in some types of aggression contexts, such as convergence or coalitionary support. Likewise, callers may aim to elicit similar behaviour from conspecifics in snake and other aggression contexts, such as ground-directed vigilance rather than escape behaviour. To test this idea, observations of receiver responses elicited from chutters and rraups emitted in different contexts, rather than only assessment of the context of emission, may be revealing. If our caller-goal theory is correct, we would predict that whether or not chutters are emitted in the context of snakes, within- or between-group aggression, they should elicit similar receiver behaviour, such as increased vigilance. Likewise, rraups elicited in different contexts, whether in eagle or aggression contexts, should nonetheless elicit similar receiver responses, such as increased cohesion. Thus, we would expect the receiver response to differ depending on the *call type*, chutters or rraups, rather than depending on the context, alarm or aggression. Playback experiments (e.g. Cheney & Seyfarth, 1988) contrasting chutters produced in predator contexts with chutters produced in aggression contexts, and likewise contrasting rraups produced in predator and non-predator contexts, could reveal whether similarity in receiver responses clusters more with respect to call type or context of production.

VII. DISCUSSION

Over the last four decades, researchers have made significant strides in understanding vocal communication from the perspective of receivers by documenting the information that individuals gain when hearing particular call types (Wheeler & Fischer, 2012). Less progress has been made in understanding vocal production in terms of why one call type is emitted instead of another, or what promotes vocal differentiation. Reviewing existing evidence, we argue that call production is driven by a combination of caller goal which determines the call type produced, and caller arousal which determines the acoustic variation observed within call types. Despite the dual contributions of caller goal and caller arousal, we propose that caller goal is the key to understanding call production. The caller-goal framework of call production offers insight into both the proximate motivations causing individuals to vocalize, and the evolution of signal differentiation. We have used examples from chimpanzees and vervets to demonstrate the potential the caller-goal framework in understanding call production, but we believe the framework is applicable to a wide range of taxa and has the potential to push forward the field of vocal communication by generating novel hypotheses.

(1) Caller goal and call function

Despite recent reconsideration of functional reference (Wheeler & Fischer, 2012), call context still dominates researchers' thinking about call production and call function (Cheney & Seyfarth, 2018). Our framework, by contrast, suggests that a significant part of hypothesis generation should derive from the calls themselves, rather than the context in which the calls are produced. If callers produce the same call type in different contexts, researchers need to entertain the hypothesis that these calls serve the same function and are produced by callers to achieve the same goal. While this may seem obvious, it is often difficult to see past the context in which a call is produced. Indeed, considering the calls as part of the same general call type may illuminate their respective functions in ways not possible when considering the calls independently because of their disparate context (as with chimpanzee 'hoos').

Viewing call type as a signal of caller goal could encourage researchers to reconsider the connections between similar calls in different contexts. Such a reframing may clarify existing communicative conundrums. For example, the proximate mechanism underlying functionally deceptive calls remains unclear (Wheeler, 2009). Considering these calls in light of a caller's goal in relation to receivers' responses, however, may clarify the situation. If the goal of a caller producing an 'avian alarm call' is to draw the attention of group mates upwards (as opposed to alerting group mates to the presence of an avian predator *per se*), producing these calls in the presence of aerial predators as well as during competitive feeding events enables callers to achieve their goals: to warn receivers of danger and to obtain food, respectively.

(2) Signal differentiation

The caller-goal framework provides a clear rationale as to when to expect call differentiation within a vocal repertoire. We would expect callers living in societies with a greater number of differentiated relationships and more distinct social roles to require more receiver responses from more receivers to achieve their goals. In such societies, natural selection should favour greater call differentiation, with more acoustically distinguishable calls, especially when living in environments where contextual information through non-vocal channels is restricted. This caller-goal centred framework of signal differentiation may potentially be able to explain differences in call repertoires between closely related species. For example, chimpanzees, gorillas, and orangutans are closely related but have radically different social organization, in which callers likely have very different responses they need to elicit from receivers. Future research should investigate whether the call repertoires in these species reflect these putative differences in caller goals.

(3) Cognition and communication

We believe our framework of call production can apply to a wide range of taxa regardless of cognition. For example,

our framework suggests that túngara frogs (*Physalaemus pustulosus*) produce mating calls in order to achieve a specific goal (to attract a mate) (Gridi-Papp, Rand, & Ryan, 2006). This view does not require that the túngara frog understands that his call attracts females – the only requirement is that the signal mostly works, and that females approach. Nevertheless, when considering theories of the trajectory of the evolution of language, the issue of cognitive input in determining caller goals becomes relevant at some point. It may be that the more a signaller understands of its social world, the more distinct goals it may be capable of or motivated to communicate. We have proposed that a caller's social structure determines the number of distinct goals a caller may have in relation to receivers. However, a caller's cognitive capacity may also have the potential to influence both the number and nature of a caller's goals. Thus, cognitive complexity (in addition to social complexity) may drive signal differentiation and communication complexity by creating new caller goals.

The caller-goal framework in the context of signal differentiation may offer insight into the evolution of language. When our ancestors developed a robust theory of mind, this capacity likely honed callers' ability to change receivers' behaviour – and minds, to achieve caller goals more efficiently. Increasing efficiency may include encoding increasingly differentiated contextual information within the call. Thus, over time, natural selection might have favoured a larger and more complex call repertoire, allowing callers to convey a greater variety of goals requiring greater diversity of receiver behaviour to achieve these goals.

VIII. CONCLUSIONS

(1) Receiver-focused interpretations of call production such as functional reference and context-specificity have been dominant in the study of mammalian vocal communication over the last 40 years. As a result, researchers have often used the contexts in which calls occur to categorize and conceptualize call types (e.g. travel calls, copulation calls, eagle alarm calls, separation calls, food calls). While a strong association between call type and context often exists, the correlation is imperfect in two ways. First, callers do not automatically produce context-specific call types when they are in the appropriate context (e.g. a vervet monkey does not produce an 'eagle alarm' every time it sees an eagle). Second, callers sometimes produce a call type in the 'wrong' context [e.g. tufted capuchin monkeys (*Cebus apella nigrinus*) produce 'alarm calls' during feeding events]. These observations suggest that particular call types do not refer to external objects in a strict sense. Some researchers have suggested that affect- or arousal-based call production can explain observed patterns of call production. However, while arousal influences call prosody (call pitch, rhythm, amplitude and tonality) within a call type, we argue that it cannot explain why individuals produce different call types.

Based on emerging evidence that calls are goal-driven, we instead argue that call type is a signal of a caller's goal. We believe that viewing call type as a signal of caller goal is the best way to understand current data on call production and will be useful for generating new research.

(2) In addition to explaining puzzling patterns of call production, a caller-goal framework is also useful for understanding the evolution of signal differentiation and call-repertoire size. Previous studies have noted links between sociality and repertoire size, but have not specified exactly why animals living in societies that are more complex might require a greater number of differentiated signals. Refocusing on the caller perspective elucidates how social complexity might favour larger call repertoires. As social complexity increases, particularly in terms of social and cooperative roles, callers may benefit from eliciting a larger number of specific receiver responses. When eliciting specific receiver responses is adaptive, we posit that specific calls that predictably elicit specific behaviour from receivers may evolve, hence requiring call diversity and larger call repertoires. We expect this to be particularly so when socially complex species live in low-visibility habitats and/or communicate over long distances.

IX. ACKNOWLEDGMENTS

We are grateful to Charles Snowdon, Andy Gersick and two anonymous reviewers for insightful comments. CC was funded by the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 679787) and the Max Planck Society.

X. REFERENCES

- BAR-ON, D. & MOORE, R. (2017). Pragmatic interpretation and signaler-receiver asymmetries in animal communication. In *Routledge Handbook for the Philosophy of Animal Minds* (eds K. ANDREWS and J. BECK). Routledge, London.
- BLUMSTEIN, D. T. (2007). The evolution, function, and meaning of marmot alarm communication. *Advances in the Study of Behavior* **37**, 371–401.
- BLUMSTEIN, D. T. & ARMITAGE, K. B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling Sciurid alarm calls. *The American Naturalist* **150**(2), 179–200.
- BRIEFER, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology* **288**(1), 1–20.
- CHENEY, D. L. & SEYFARTH, R. M. (1985). Vervet monkey alarm calls: manipulation through shared information? *Behaviour* **94**(1), 150–166.
- CHENEY, D. L. & SEYFARTH, R. M. (1988). Vervet monkey assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour* **36**(2), 477–486.
- CHENEY, D. L. & SEYFARTH, R. M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. University of Chicago Press, Chicago.
- CHENEY, D. L. & SEYFARTH, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences* **115**(9), 1774–1779.
- CROSS, R. G., MCCOWAN, B. & RAMAKRISHNAN, U. (2007). Threat-related acoustical differences in alarm calls by wild bonnet macaques (*Macaca radiata*) elicited by python and leopard models. *Ethology* **113**(4), 352–367.
- CROCKFORD, C. & BOESCH, C. (2003). Context-specific calls in wild chimpanzees, *Pan troglodytes verus*: analysis of barks. *Animal Behaviour* **66**(1), 115–125.
- CROCKFORD, C. & BOESCH, C. (2005). Call combinations in wild chimpanzees. *Behaviour* **142**(4), 397–421.

- CROCKFORD, C., GRUBER, T. & ZUBERBÜHLER, K. (2018). Chimpanzee quiet hoo variants differ according to context. *Royal Society Open Science* **5**(5), 172066.
- CROCKFORD, C., WITTIG, R. M., MUNDRY, R. & ZUBERBÜHLER, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology* **22**, 142–146.
- CROCKFORD, C., WITTIG, R. M. & ZUBERBÜHLER, K. (2015). An intentional vocalization draws others' attention: a playback experiment with wild chimpanzees. *Animal Cognition* **18**(3), 581–591.
- CROCKFORD, C., WITTIG, R. M. & ZUBERBÜHLER, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances* **3**(11), e1701742.
- CRYSTAL, D. (2010). *The Cambridge Encyclopedia of Language*, 3rd Edition. Cambridge University Press, Cambridge.
- CUTLER, A., DAHAN, D. & VAN DONSELAAR, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech* **40**(Pt 2), 141–201.
- DENNETT, D. C. (1983). Intentional systems in cognitive ethology: the “Panglossian paradigm” defended. *Behavioral and Brain Sciences* **6**(3), 343–355.
- DICKINSON, A. & BALLEINE, B. W. (2000). Causal cognition and goal-directed action. In *The Evolution of Cognition* (eds C. HEYES & L. HUBER). MIT Press, Cambridge.
- DUFFY, K. G., WRANGHAM, R. W. & SILK, J. B. (2007). Male chimpanzees exchange political support for mating opportunities. *Current Biology* **17**(15), R586–R587.
- ESCH, H. C., SAYIGH, L. S., BLUM, J. E. & WELLS, R. S. (2009). Whistles as potential indicators of stress in bottlenose dolphins (*Tursiops truncatus*). *Journal of Mammalogy* **90**(3), 638–650.
- FALLON, B. L., NEUMANN, C., BYRNE, R. W. & ZUBERBÜHLER, K. (2016). Female chimpanzees adjust copulation calls according to reproductive status and level of female competition. *Animal Behaviour* **113**, 87–92.
- FICHTEL, C. & HAMMERSCHMIDT, K. (2002). Responses of redfronted lemurs to experimentally modified alarm calls: evidence for urgency-based changes in call structure. *Ethology* **108**(9), 763–778.
- FICHTEL, C. & MANSER, M. (2010). Vocal communication in social groups. In *Animal Behaviour: Evolution and Mechanisms* (ed. P. KAPPELER), pp. 29–54. Springer Berlin Heidelberg, Heidelberg.
- FICHTEL, C., HAMMERSCHMIDT, K. & JÜRGENS, U. (2001). On the vocal expression of emotion: a multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour* **138**(1), 97–116.
- FISCHER, J. (2017). Primate vocal production and the riddle of language evolution. *Psychonomic Bulletin & Review* **24**(1), 72–78.
- FISCHER, J. & PRICE, T. (2016). Meaning, intention, and inference in primate vocal communication. *Neuroscience & Biobehavioral Reviews* **82**, 22–31.
- FISCHER, J., HAMMERSCHMIDT, K., CHENEY, D. L. & SEYFARTH, R. M. (2001). Acoustic features of female *Chacma* baboon barks. *Ethology* **107**(1), 33–54.
- FISCHER, J., HAMMERSCHMIDT, K. & TODT, D. (1995). Factors affecting acoustic variation in barbary macaque (*Macaca sylvanus*) disturbance calls. *Ethology* **101**(1), 51–66.
- FURRER, R. D. & MANSER, M. B. (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *The American Naturalist* **173**(3), 400–410.
- GENTY, E., CLAY, Z., HOBATTER, C. & ZUBERBÜHLER, K. (2014). Multi-modal use of a socially directed call in bonobos. *PLoS One* **9**, e84738.
- GERSICK, A. S., CHENEY, D. L., SCHNEIDER, J. M., SEYFARTH, R. M. & HOLEKAMP, K. E. (2015). Long-distance communication facilitates cooperation among wild spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour* **103**, 107–116.
- GOODALL, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Belknap Press of Harvard University Press, Cambridge.
- GOUZOULES, H. & GOUZOULES, S. (1989). Design features and developmental modification of pigtail macaque, *Macaca nemestrina*, agonistic screams. *Animal Behaviour* **37**, 383–401.
- GRIDI-PAPP, M., RAND, A. S. & RYAN, M. J. (2006). Animal communication: complex call production in the túngara frog. *Nature* **441**(7089), 38.
- GRUBER, T. & GRANDJEAN, D. (2017). A comparative neurological approach to emotional expressions in primate vocalizations. *Neuroscience & Biobehavioral Reviews* **73**, 182–190.
- GRUBER, T. & ZUBERBÜHLER, K. (2013). Vocal recruitment for joint travel in wild chimpanzees. *PLoS One* **8**(9), e76073.
- GUSTISON, M. L., LE ROUX, A. & BERGMAN, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**(1597), 1847–1859.
- HIRST, D. & DI CRISTO, A. (1998). *Intonation Systems: A Survey of twenty Languages*. Cambridge University Press, Cambridge.
- KAJIKAWA, S. & HASEGAWA, T. (2000). Acoustic variation of pant hoot calls by male chimpanzees: a playback experiment. *Journal of Ethology* **18**(2), 133–139.
- KALAN, A. K. & BOESCH, C. (2015). Audience effects in chimpanzee food calls and their potential for recruiting others. *Behavioral Ecology and Sociobiology* **69**(10), 1701–1712.
- KARAKASHAIN, S. J., GYGER, M. & MARLER, P. (1988). Audience effects on alarm calling in chickens (*Gallus gallus*). *Journal of Comparative Psychology* **102**(2), 129–135.
- KREBS, J. R. & DAWKINS, R. (1984). Animal signals: mind-reading and manipulation. In *Behavioural Ecology an Evolutionary Approach*, 2nd Edition (eds J. R. KREBS and N. B. DAVIES). Blackwell Science Ltd, Hoboken.
- LAIDRE, M. E. & JOHNSTONE, R. A. (2013). Animal signals. *Current Biology* **23**, R829–R833.
- LANGERGRABER, K. E., MITANI, J. C. & VIGILANT, L. (2007). The limited impact of kinship on cooperation in wild chimpanzees. *Proceedings of the National Academy of Sciences* **104**(19), 7786–7790.
- MACEDONIA, J. M. & EVANS, C. S. (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* **93**, 177–197.
- MANSER, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**(1483), 2315–2324.
- MANSER, M. B. (2010). The generation of functionally referential and motivational vocal signals in mammals. *Handbook of Behavioural Neuroscience* **19**, 477–486.
- MAYR, E. (1961). Cause and effect in biology. *Science* **134**, 1501–1506.
- MCCOMB, K. & SEMPLE, S. (2005). Coevolution of vocal communication and sociality in primates. *Biological Letters* **1**(4), 381–385.
- MITANI, J. C. & NISHIDA, T. (1993). Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour* **45**(4), 735–746.
- MORTON, E. (2017). *Animal Vocal Communication*. Cambridge University Press, Cambridge.
- OWINGS, D. & MORTON, E. (1998). *Animal Vocal Communication: A New Approach*. Cambridge University Press, Cambridge.
- OWREN, M. J. & RENDALL, D. (2001). Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology Issues, News, and Reviews* **10**(2), 58–71.
- OWREN, M. J., AMOSS, R. T. & RENDALL, D. (2011). Two organizing principles of vocal production: implications for nonhuman and human primates. *American Journal of Primatology* **73**(6), 530–544.
- PAPWORTH, S., BÖSE, A. S., BARKER, J., SCHEL, A. M. & ZUBERBÜHLER, K. (2008). Male blue monkeys alarm call in response to danger experienced by others. *Biological Letters* **4**(5), 472–475.
- POLLARD, K. A. & BLUMSTEIN, D. T. (2012). Evolving communicative complexity: insights from rodents and beyond. *Philosophical Transactions of the Royal Society B* **367**(1597), 1869–1878.
- PREMACK, D. (1972). Concordant preferences as a precondition for affective but not for symbolic communication (or how to do experimental anthropology). *Cognition* **1**(2), 251–264.
- 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**, srep13220.
- RENDALL, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *The Journal of the Acoustical Society of America* **113**(6), 3390–3402.
- RENDALL, D., OWREN, M. J. & RYAN, M. J. (2009). What do animal signals mean? *Animal Behaviour* **78**(2), 233–240.
- RENDALL, D., SEYFARTH, R. M., CHENEY, D. L. & OWREN, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour* **57**, 583–592.
- RUBOW, J., CHERRY, M. I. & SHARPE, L. L. (2017). Calling for help: dwarf mongoose recruitment calls inform receivers about context and elicit disparate responses. *Animal Behaviour* **127**, 7–14.
- RUIZ, R., ABSIL, E., HARMEGINIES, B., LEGROS, C. & POCH, D. (1996). Time- and spectrum-related variabilities in stressed speech under laboratory and real conditions. *Speech Communication* **20**(1–2), 111–129.
- 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.
- SCHAMBERG, I., CHENEY, D. L., CLAY, Z., HOHMANN, G. & SEYFARTH, R. M. (2017). Bonobos use call combinations to facilitate inter-party travel recruitment. *Behavioural Ecology and Sociobiology* **71**(4), 75.
- SCHEHKA, S., ESSER, K. H. & ZIMMERMANN, E. (2007). Acoustical expression of arousal in conflict situations in tree shrews (*Tupaia belangeri*). *Journal of Comparative Physiology* **193**(8), 845–852.
- SCHEL, A. M., MACHANDA, Z., TOWNSEND, S. W., ZUBERBÜHLER, K. & SLOCOMBE, K. E. (2013a). Chimpanzee food calls are directed at specific individuals. *Animal Behaviour* **86**(5), 955–965.
- SCHEL, A. M., MACHANDA, Z., TOWNSEND, S. W., ZUBERBÜHLER, K. & SLOCOMBE, K. E. (2013b). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS One* **8**(10), e76674.
- SCHRADER, L. & TODT, D. (1993). Contact call parameters covary with social context in common marmosets, *Callithrix j. jacchus*. *Animal Behaviour* **46**(5), 1026–1028.
- SCHÜLKE, O., BHAGAVATULA, J., VIGILANT, L. & OSTNER, J. (2010). Social bonds enhance reproductive success in male macaques. *Current Biology* **20**(24), 2207–2210.
- SEYFARTH, R. M. & CHENEY, D. L. (1997). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning & Behavior* **25**(3), 249–267.
- SEYFARTH, R. M. & CHENEY, D. L. (2018). Pragmatic flexibility in primate vocal production. *Current Opinion in Behavioral Sciences* **21**, 56–61.
- 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.

- 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.
- SILK, J. B., ALBERTS, S. C. & ALTMANN, J. (2003). Social bonds of female baboons enhance infant survival. *Science* **302**(5648), 1231–1234.
- SILK, J. B., BEEHNER, J. C., BERGMAN, T., CROCKFORD, C., ENGH, A. L., MOSCOVICE, L. R., WITTIG, R. M., SEYFARTH, R. M. & CHENEY, D. L. (2010). Female chacma baboons form strong, equitable, and enduring social bonds. *Behavioral Ecology and Sociobiology* **64**, 1733–1747.
- SILK, J. B., SEYFARTH, R. M. & CHENEY, D. L. (2016). Strategic use of affiliative vocalizations by wild female baboons. *PLoS One* **11**(10), e0163978.
- SIMONS, G. F. & FENNIG, C. D. (eds) (2018). *Ethnologue: Languages of the World, Twenty*, 1st Edition (). SIL International, Dallas.
- SLOCOMBE, K. E. & ZUBERBÜHLER, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences* **104**(43), 17228–17233.
- SMITH, C. L., TAYLOR, A. & EVANS, C. S. (2011). Tactical multimodal signalling in birds: facultative variation in signal modality reveals sensitivity to social costs. *Animal Behaviour* **82**(3), 521–527.
- SNOWDON, C. T. (1990). Language capacities of nonhuman animals. *American Journal of Physical Anthropology* **33**(S11), 215–243.
- SNOWDON, C. T. & DE LA TORRE, S. (2002). Multiple environmental contexts and communication in pygmy marmosets (*Cebuella pygmaea*). *Journal of Comparative Psychology* **116**(2), 182–188.
- SOLTIS, J., LEIGHTY, K. A., WESOLEK, C. M. & SAVAGE, A. (2009). The expression of affect in African elephant (*Loxodonta africana*) rumble vocalizations. *Journal of Comparative Psychology* **123**(2), 222–225.
- STRUSAKER, T. T. (1967). Ecology of Vervet monkeys (*Cercopithecus Aethiops*) in the Masai-Amboseli game reserve, Kenya. *Ecology* **48**(6), 891–904.
- TOWNSEND, S. W. & MANSER, M. B. (2013). Functionally referential communication in mammals: the past, present and the future. *Ethology* **119**(1), 1–11.
- TOWNSEND, S. W., DESCHNER, T. & ZUBERBÜHLER, K. (2008). Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS One* **3**(6), e2431.
- TOWNSEND, S. W., KOSKI, S. E., BYRNE, R. W., SLOCOMBE, K. E., BICKEL, B., BOESCHLE, M., BRAGA GONCALVES, I., BURKHAR, J. M., FLOWER, T., GAUNET, F., GLOCK, H. J., GRUBER, T., JANSEN, D. A. W. A. M., LIEBAL, K., LINKE, A., et al. (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews* **92**, 1427–1433.
- TRIPOVICH, J. S., CANFIELD, R., ROGERS, T. L. & ARNOULD, J. P. Y. (2008). Characterization of Australian fur seal vocalizations during the breeding season. *Marine Mammal Science* **24**, 913–928.
- WHEELER, B. C. (2009). Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society of London B: Biological Sciences* **276**(1669), 3013–3018.
- WHEELER, B. C. & FISCHER, J. (2012). Functionally referential signals: a promising paradigm whose time has passed. *Evolutionary Anthropology* **21**(5), 195–205.
- WICH, S. A. & DE VRIES, H. (2006). Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society of London B: Biological Sciences* **273**(1587), 735–740.
- WITTIG, R. M., CROCKFORD, C., WELTRING, A., LANGERGRABER, K. E., DESCHNER, T. & ZUBERBÜHLER, K. (2016). Social support reduces stress hormone levels in wild chimpanzees across stressful events and everyday affiliations. *Nature Communications* **7**, 13361.

(Received 2 November 2017; revised 22 May 2018; accepted 23 May 2018)