



Call combinations, vocal exchanges and interparty movement in wild bonobos



Isaac Schamberg^{a,*}, Dorothy L. Cheney^b, Zanna Clay^c, Gottfried Hohmann^d,
Robert M. Seyfarth^a

^a Department of Psychology, University of Pennsylvania, Philadelphia, PA, U.S.A.

^b Department of Biology, University of Pennsylvania, Philadelphia, PA, U.S.A.

^c Department of Psychology, University of Birmingham, Birmingham, U.K.

^d Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

ARTICLE INFO

Article history:

Received 1 April 2016

Initial acceptance 17 May 2016

Final acceptance 27 July 2016

MS. number: A16-00288R

Keywords:

bonobo
call combination
call exchange
fission–fusion
persistence
vocalization

The vocal repertoire of nonhuman primates is largely fixed. Individuals produce their species-specific vocalizations from a young age, and do not acquire new call types over their lifetime. Despite these limitations, however, monkeys and apes are able to increase their vocal flexibility in several ways, including subtle acoustic modification, call combinations, turn-taking and call persistence. Although primates have been observed to utilize these communicative features, the extent to which they integrate these abilities is not known. Here we show that certain long-distance calls produced by wild bonobos, *Pan paniscus*, assimilate several aspects of vocal flexibility in ways not previously documented in nonhuman primates. Communication between foraging parties exhibits context-specific call combinations relating to the movement of caller, call modifications that potentially target particular individuals, call-and-answer exchanges in which the initial caller's behaviour depends on the listener's reply, and possible persistence in call production. The selective pressure exerted by bonobos' fission–fusion social structure has likely favoured the integration of these communicative capabilities.

© 2016 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

In recent years, field studies of monkeys and apes have drawn attention to the importance of call combinations in primate vocal communication. The extensive use of call combinations by certain species has led to some reconsideration of previous assumptions about the inflexible nature of primate vocal production. While the number of distinct call types a species can produce appears to be fixed, callers can increase their effective vocal repertoire through the use of call combinations (reviewed in Zuberbühler & Lemasson, 2013).

Multiple species provide evidence that the information contained in call combinations differs from the sum of the information contained in the individual call types. Male Campbell's monkeys, *Cercopithecus campbelli*, for example, use six different call types to produce nonrandom call sequences in response to predators and other environmental disturbances. Two features of their calling system have invited particular interest. First, 'boom' calls at the

beginning of call sequences seem to alter the meaning of subsequent calls in the sequence. Males produce sequences of 'krak-oo' calls in response to the calls of leopards, *Panthera pardus*, and crowned-hawk eagles, *Stephanoaetus coronatus*, and after hearing other monkeys' alarm calls towards those predators. In response to environmental disturbances such as falling branches, however, callers will produce a 'boom' call before uttering a sequence of 'krak-oo' calls. Heterospecific listeners react with predator avoidance behaviour to 'krak-oo' sequences, but largely ignore 'boom-krak-oo' sequences (Zuberbühler, 2002). Second, the composition of different call types correlates both with predator type (ground or aerial) and mode of detection (visual or auditory) (Ouattara, Lemasson, & Zuberbühler, 2009a, b).

Another example of call combinations comes from putty-nosed monkeys, *Cercopithecus nictitans*, which produce 'pyow' calls in response to general environmental disturbances as well as to leopards, and distinct 'hack' calls primarily in response to crowned eagles. When the two call types are produced sequentially in 'pyow-hack' sequences, however, the calls are associated with group movement that appear unrelated to leopards or eagles (Arnold & Zuberbühler, 2006a, b, 2008, 2012).

* Correspondence: I. Schamberg, Department of Psychology, University of Pennsylvania, Philadelphia, PA, U.S.A.

E-mail address: sisaac@sas.upenn.edu (I. Schamberg).

The use of call combinations is not limited to forest guenons. Gibbons (Hylobatidae), black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*), and titi monkeys (*Callicebus nigrifrons*) also produce call sequences in which the composition of call types corresponds to predator presence, predator type and/or predator location (gibbons: Clarke, Reichard, & Zuberbühler, 2006; black-and-white colobus: Schel, Tranquilli, & Zuberbühler, 2009; titi monkeys: Cäsar, Zuberbühler, Young, & Byrne, 2013).

Call combinations also occur in nonpredator contexts. Bonobos, *Pan paniscus*, produce call sequences containing five different call types when feeding near other individuals. The proportion of each call type in a sequence correlates with feeding preferences (or possibly food quality), and listeners can use this information to guide their own foraging behaviour (Clay & Zuberbühler, 2009, 2011). Chimpanzees, *Pan troglodytes*, too, appear to use call combinations extensively in social contexts such as feeding and travelling. Nearly 50% of their vocalizations are produced as part of a call combination, but the function of these combinations remains ambiguous (Crockford & Boesch, 2005). Similarly, female Diana monkeys, *Cercopithecus diana*, use four call types to produce call combinations during social activities; again, however, the function of these call combinations is not yet known (Candiotti, Zuberbühler, & Lemasson, 2012).

Taken together, these observations suggest that the information contained in call combinations may not derive, in a straightforward way, from the information contained in their constituent calls. How the informational content of call combinations arises (and differs) from individual call types is an active topic of research. It remains unclear whether any of the examples of call combinations are compositional, in that the 'meaning' of the combination is based on the 'meanings' of its constituent calls, or if the calls combine in a noncompositional manner to convey information unrelated to the informational content of the constituent calls (Schlenker et al., 2014; Schlenker, Chemla, Arnold, & Zuberbühler, 2016).

In order to make progress toward understanding how primates combine call types, it is critical to collect data across populations and taxa. Comparison of different populations of Campbell's monkeys has already led to new testable hypotheses about their alarm call system (Schlenker et al., 2014, 2016), and reviews across taxa have generated new hypotheses about the evolution of language (Collier, Bickel, van Schaik, Manser, & Townsend, 2014). Further comparative work has the potential to clarify the relationship between the selective pressures acting on a species' communication systems and the call combinations exhibited by those same species.

Species exhibiting fission–fusion dynamics, in which members of a single social group regularly divide into smaller social units, may be of particular interest for investigations into call combinations. The demands placed on individuals living in fission–fusion societies may place selection pressures on individuals' cognitive and communicative abilities because animals must maintain relationships with individuals whom they may encounter irregularly, and coordinate their behaviour and movement with out-of-sight individuals (Aureli et al., 2008). Several studies have documented the role of interparty communication in fission–fusion societies. Bonobos and spider monkeys (*Ateles geoffroyi*), for example, both use long-distance vocalizations facultatively to maintain contact and coordinate movement with out-of-sight individuals (bonobos: Hohmann & Fruth, 1994; White, Waller, Boose, Merrill, & Wood, 2015; spider monkeys: Ramos-Fernández, 2005; Spehar & Di Fiore, 2013). Chimpanzees also use vocalizations to coordinate with out-of-sight individuals (Fedurek, Donnellan, & Slocombe, 2014), and they appear to modify their call production based on knowledge of which individuals are nearby (Kalan & Boesch, 2015; Mitani & Nishida, 1993). Hyenas (*Crocuta crocuta*), in addition to exchanging long-distance vocalizations between out-of-site

individuals, produce context-related call subtypes that facilitate mutually beneficial movement patterns between callers and receivers (Gersick, Cheney, Schneider, Seyfarth, & Holekamp, 2015).

While not all vocalizations exchanged between parties in fission–fusion societies involve call combinations, the uncertainty inherent in movement patterns of fission–fusion societies (whether individual A approaches individual B or vice versa) may create situations in which contact calls that only provide information about identity and location are insufficient to facilitate effective group movement. Indeed, previous studies on interparty communication have indicated that such calls do not signal the relative direction of interparty movement (Fedurek et al., 2014; Spehar & Di Fiore, 2013). Call combinations are one mechanism by which individuals can convey additional information about context or caller motivation to listeners and, thus, potentially, reduce the uncertainty involved in interparty movement. Given the documented use of call combinations in chimpanzees and bonobos, call sequences probably play a role in interparty communication in these two species.

Here we present data on the use of long-distance vocalizations by bonobos during interparty movement. Bonobos produce several signals during the context of interparty movement and combine these signals nonrandomly. Here, we focus on two call types: the high hoot (HH) and the whistle–high hoot combination (W + HH). We report that wild bonobos produce the W + HH call combination when apparently highly motivated to move from one foraging party to another. Callers are significantly more likely to move to a new party after producing a W + HH combination than after producing HHs alone, especially if the caller receives a response from the group it is about to join. Callers who do not receive a response often call again, underscoring their motivation to receive a response before joining another party. Callers also modify the acoustic structure of their combined calls in a manner that distinguishes between those given spontaneously and those given in response to another call. We suggest that these four features have adaptive value because they coordinate activity with out-of-sight group-mates.

METHODS

Data Collection

Bonobos form long-term, stable communities in which all members share a home range and form an exclusive reproductive unit (Kano, 1992). Within a community, individuals form temporary subgroups, or 'parties', that travel and forage separately from other parties. Parties are unpredictable in size (ranging from one individual to the entire community), duration (lasting from several minutes to several days) and composition (because animals do not always form a party with the same individuals).

For 13 months between July 2011 and March 2014, we sampled behaviour and recorded vocalizations from 18 free-ranging adults (7 males and 11 females) at the LuiKotale field site in the Bandundu province of the Democratic Republic of Congo (Hohmann & Fruth, 2003). Data collection included focal animal sampling, ad libitum sampling and scan sampling (Altmann, 1974). Data on rates of vocalizations were calculated from focal sampling; all other analyses use both focal and ad libitum sampling. We obtained 1224 h of ad libitum sampling and 117 h of 15 min focal animal sampling. No subject was sampled within 1 h of its last focal sample and effort was made to sample each subject in a party once before sampling any animal a second time. Focal samples included continuous data on vocal behaviour and the occurrence of fissions or fusions in the focal animal's party. Observers also collected data on affiliative, agonistic and feeding behaviour, as well as dominance interactions. Focal observations were supplemented by ad libitum observations

of the same behaviours and vocalizations. Finally, every 15 min observers conducted a party composition scan in which the identity of all bonobos visible was recorded. Party composition was defined as all individuals visible to observers or known to be within a radius of 50 m of the focal animal (Lehmann & Boesch, 2004). Observers visually scanned the surrounding area and conferred with other observers in order to identify all animals in the party. Scans also included currently out-of-sight bonobos that were known to be present based on observations in the previous 15 min. These fixed-time party composition scans provided the data for calculating baseline changes in party composition.

In addition to behavioural data, observers made continuous audio recordings of all directly observed vocalizations. Recordings were made using a PMD660 Marantz digital recorder and a Sennheiser ME66 microphone at sample rate of either 44.1 kHz or 48 kHz.

Spectrograms of audio files were created with WaveSurfer (version 1.88p, <https://wavesurfer-js.org/>) and RavenPro (version 1.5, <http://www.birds.cornell.edu/brp/raven/ravenversions>). Call types and call combinations were visually distinguishable in spectrograms (see Results, Figs. 1 and 3). Classification of call types followed descriptions of the bonobo vocal repertoire in captivity (de Waal, 1988) and in the wild (Bermejo & Omedes, 1999).

When an individual produced a vocalization, observers noted the call type, the activity of the caller, the identity of individuals within 10 m of the caller, immediate behavioural change after the

call, and all vocalizations produced by the caller and by other individuals that preceded or followed the call. A caller was considered to have approached another party, if within 15 min after producing the call, it travelled more than 50 m and encountered individuals that were not part of its most recent party. We chose 15 min as our time limit based on personal observations of typical travel time between parties and the length of focal animal samples.

Observers categorized each vocalization produced by subjects as a 'spontaneous' or 'response' call. 'Spontaneous' calls were those given in the absence of any calls by individuals outside the subject's party during the 30 s prior to the focal animal's call. 'Response' calls were those produced within 10 s of vocalizations from another party. Observers also noted whether each call received a 'response', that is, whether it was followed within 10 s by vocalizations from bonobos outside the subject's party. We chose 10 s as the window for response vocalizations because, based on previous observations, bonobos occasionally take several seconds to respond to vocalizations. In most cases, however, responses were produced immediately after the spontaneous calls.

Because of the fragmented and unpredictable nature of bonobos' parties, observers were unable to obtain simultaneous audio recording from both the spontaneous caller and the response caller during a single call exchange. That is, the observer either recorded the initial, spontaneous calls and then heard the response from another party, or recorded the response calls just after hearing spontaneous calls from another party. Although observers could clearly hear vocalizations from other parties, only calls that were audiorecorded were used in analyses. Vocalizations produced by individuals in other parties were only used to classify audiorecorded vocalizations as either spontaneous or response calls.

Data Sets

To answer our different questions, we used overlapping but nonidentical data sets. To examine whether callers subsequently approached and joined another party, we used observations of 50 W + HH combinations (34 spontaneous calls and 16 response calls) recorded from seven adult males and seven adult females. We compared these W + HHs combinations to 75 observations of HHs alone (44 spontaneous calls and 31 response calls) recorded from seven adult males and eight adult females for which the caller's subsequent movement was definitively known. Some of these observations occurred as part of larger communicative events (i.e. callers had produced multiple bouts of HHs or W + HHs within a 10 min window). To maintain independence between observations, we only included the final HHs or W + HHs given by a caller during a communicative event.

For the analysis on persistence in call production, we included all observed W + HHs (72 spontaneous W + HHs and 16 response W + HHs) recorded from seven adult males and seven adult females. We also examined call persistence in all the HHs produced in the first 5 min of focal animal samples (31 spontaneous HHs and 23 response HHs) recorded from seven adult males and 10 adult females.

For the acoustic analysis of whistle types, we included all W + HHs for which the audio recording was of a high enough quality to conduct the appropriate classification (56 spontaneous W + HHs and 13 response W + HHs) recorded from six adult males and two adult females.

Statistical Analysis

To test whether certain calls and call combinations were followed by different behaviours by callers, listeners, or both, we used generalized mixed models ('glmer' function in 'lmerTest' package' in R version 3.1.2 GUI 1.65 Snow Leopard build (6833)). Because

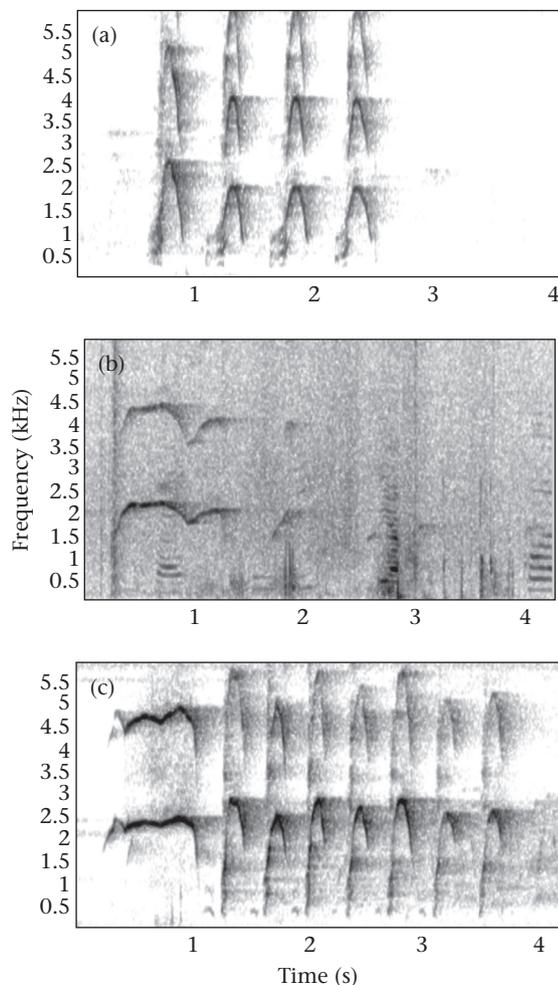


Figure 1. Spectrograms of different call types: (a) high hoot bout containing four call units, (b) whistle and (c) whistle–high hoot combination.

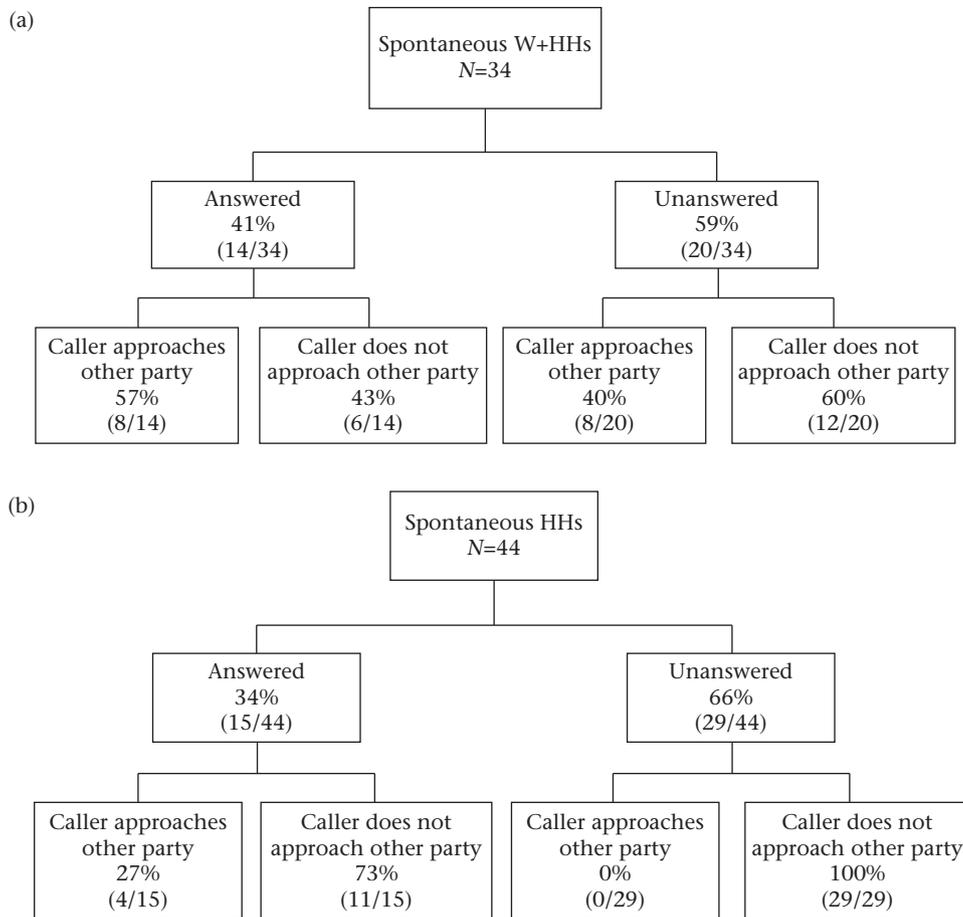


Figure 2. The outcomes of observed (a) spontaneous whistle–high hoot combinations (W + HHs) and (b) spontaneous high hoots (HHs). Data on W+HHs were collected from seven adult males and seven adult females. Data on HHs alone were collected from seven adult males and eight adult females. Data are based on both focal and ad libitum observations.

different individual callers contributed in different proportions to our pooled data, we entered caller ID as a random factor.

Ethical Note

Subjects for this study were 18 free-ranging adult bonobos. Data collection consisted only of behavioural observations. The study was conducted in accordance with the current laws in the United States, Germany and the Democratic Republic of the Congo. The research was approved by the Animal Care and Use Committee of the University of Pennsylvania (Protocol no. 804117).

RESULTS

High Hoots and Whistle–High Hoot Combinations

A common vocalization among bonobos is the high hoot (HH) (Fig. 1a), a loud, tonal call (de Waal, 1988) given in a variety of nonaggressive contexts, and occurring in bouts consisting of 1–27 acoustic units (Hohmann & Fruth, 1994), each with an inverted U-shaped frequency contour. High hoots are audible for at least 700 m in the forest (I. Schamberg, personal observation). They appear to be individually distinctive, and previous research suggests that they may facilitate the joining of separated parties (Hohmann & Fruth, 1994; White et al., 2015). In our study, bonobos produced bouts composed exclusively of HHs at an overall mean \pm SD rate of 1.09 ± 0.46 bouts per individual per hour (male rate = 1.26 ± 0.35 , female rate = 0.98 ± 0.50).

High hoots are also produced in combination with long, tonal vocalizations, termed ‘whistles’ (Bermejo & Omedes, 1999) (Fig. 1b). Whistles were almost always produced as the initial call of a call combination, either with high hoots or ‘contest hoots’, an agonistic vocalization (Genty, Clay, Hobaiter, & Zuberbühler, 2014). Individuals also occasionally produced whistles as a stand-alone call in the absence of either high hoots or contest hoots, corroborating previous research identifying the whistles as a distinct call type (Bermejo & Omedes, 1999).

Whistle–high hoot combinations (W + HHs, Fig. 1c) consisted of one or two whistles and between one and 13 HH units. Whistles always preceded HHs. Fourteen of 18 subjects produced at least one W + HH combination (Supplementary Table S1). Males appeared to produce W + HHs more frequently than females. In our primary data set, males produced 74% (37/50) of W + HHs. Furthermore, high-ranking males appeared to be less likely to produce W + HHs than mid- or low-ranking males. The two highest-ranking males produced 11% (4/37) of the W + HHs produced by males, while the other five males produced 89% (33/37) of the calls. Overall, subjects produced bouts containing W + HHs at an overall rate of 0.11 ± 0.25 bouts per individual per hour.

Effect of Call Type and Call Exchanges on Post-call Behaviour

Our observations indicated that both HHs and W + HHs were associated with interparty movement. During 468 focal animal samples, subjects approached and joined another party in 8% (36/468) of samples. Of these approaches, 58% were preceded by HHs,

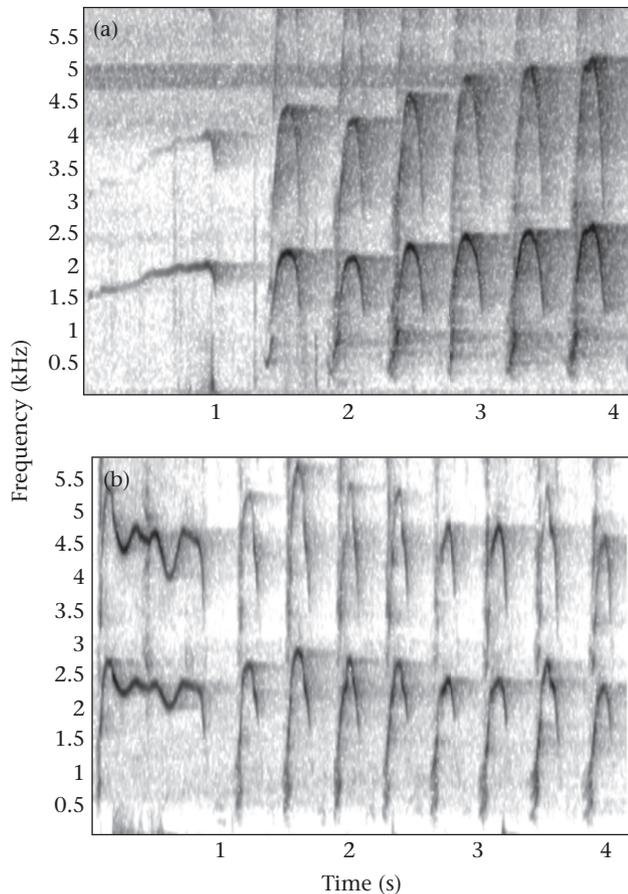


Figure 3. Spectrograms of (a) flat and (b) squiggle whistle–high hoot combinations (W + HHs).

11% were preceded by W + HHs, and 31% occurred in the absence of any long-distance calls (HHs or W + HHs). Thus the majority (69%) of interparty movement events were preceded by long-distance vocalizations.

Whether or not a caller produced HHs in combination with a whistle appeared to be influenced both by the caller's motivation to approach and join another party and by whether or not the caller's vocalizations were part of a call exchange (Fig. 2). To examine the effects of call type and the occurrence of a response on callers' approach behaviour, we ran a generalized linear mixed model (GLMM) with a logistic link function. We used a binomial outcome variable (approaching/not approaching another party) as the dependent measure, call type (HHs or W + HHs) and the presence or absence of a responding vocalization (with one exception, always a HH alone) as predictor variables, and caller ID as a random factor. We used likelihood ratio test to compare a full model that included all predictors against a null model that included only the random effects (Crawley, 2014). The full model fitted the data significantly better than the null model that included only ID as a random effect ($\chi^2_2 = 20.42$, $P = 0.000$). Given this result, we ran a single model with two binomial predictors (whistle/no whistle, response/no response). Both predictors were significant (whistle/no whistle: $\beta = 2.6$, $SE = 0.9$, $z = 3.0$, $P = 0.003$; response/no response: $\beta = 2.0$, $SE = 0.9$, $z = 2.2$, $P = 0.025$).

In summary, callers were more likely to approach and join another party after producing W + HHs than after producing HHs only, and more likely to approach after receiving a response. Callers were most likely to approach after both producing W + HHs and receiving a response (Fig. 2).

We also examined the relationship between receiving an apparent response and subsequent call production. Callers that produced spontaneous W + HHs and received an apparent response tended to be less likely to produce additional W + HHs within 10 min of the initial call bout than callers that did not receive an apparent response (GLMM: $\beta = -1.1$, $SE = 0.6$, $z = -1.8$, $P = 0.070$). Production of subsequent HHs was unrelated to whether a caller received a response (GLMM: $\beta = 0.1$, $SE = 1.5$, $z = 0.0$, $P = 0.96$).

The preceding results were derived from instances in which the observer recorded vocalizations from the individual who initiated the calling bout, and then noted whether there was a vocal response from an (unidentified) caller in another party. There were also cases, however, when the observer recorded the responder's calls after hearing spontaneous calls from an (unidentified) caller in another party. When we included these cases in our analysis of approach behaviour, sample size increased from 44 HHs and 34 W + HHs to 75 HHs and 50 W + HHs and the correlation between producing W + HHs and approaching another party became stronger (likelihood ratio test: $\chi^2_2 = 27.679$, $P = 0.000$; GLMM: whistle/no whistle: $\beta = 2.1$, $SE = 0.5$, $z = 4.4$, $P = 0.000$, exchange/no exchange: $\beta = 2.2$, $SE = 0.6$, $z = 3.6$, $P = 0.000$).

Similarly, inclusion of response calls in the analysis on persistence strengthened the trend in the data for spontaneous calls only. Specifically, callers who produced W + HHs that were part of a call exchange produced additional W + HHs in 22% (8/35) of cases, whereas callers who produced W + HHs that were not part of a call exchange produced W + HHs again in 47% (25/53) of cases (GLMM: $\beta = -1.0$, $SE = 0.5$, $z = -1.9$, $P = 0.052$). In other words, when callers produced W + HHs that did not receive a response they often gave additional W + HHs, whereas callers that produced HHs were equally likely to produce additional HHs whether they received a response or not.

Effect of Party Composition and Context on Post-call Behaviour

Social factors, such as party size and number of females in a party, were unrelated to callers' behaviour. The mean \pm SD party size when caller's subsequently approached another party was 6.1 ± 4.9 ($N = 13$); party size when the caller did not subsequently approach another party was 6.4 ± 3.2 ($N = 32$). The number of females in the caller's party when the caller subsequently approached another party was 3.9 ± 3.3 ($N = 13$); the number of females in the caller's party when the caller did not subsequently approach another party was 4.2 ± 2.1 ($N = 32$). Neither party size nor number of females in the party was a significant predictor in a GLMM with approaching/not approaching another party as the dependent measure and ID as a random factor (party size: $\beta = 0.2$, $SE = 0.3$, $z = 0.2$, $P = 0.822$; number of females in party: $\beta = -0.1$, $SE = 0.5$, $z = -0.3$, $P = 0.757$).

The different outcomes of HHs and W+HHs could have been a by-product of the different contexts in which bonobos produced these two call types. The context of production for HHs and W+HHs, however, were very similar. Of the 131 HH-only bouts produced by focal animals, 51% (67/131) were given during feeding, 29% (38/131) during periods of rest or grooming and 20% (26/131) while travelling. Of the W + HH bouts with unambiguous contexts, 38% (14/37) were given during feeding, 43% (16/37) during periods of rest and 19% (7/37) while travelling. After approaching and joining another party, callers fed in 60% (12/20) of instances in which the context was unambiguous, travelled in 40% (6/20) of instances and rested in 10% (2/20) of instances.

Production of W + HHs did not seem to differ according to the distance separating parties. We had precise measurements of the distance between the location of the call and the location of the

subsequent fusion for 10 W + HH events. The mean \pm SD distance was 207 ± 140 m with a range of distances between 59 m and 536 m. Thus, W + HHs occurred at a wide range of interparty distances, making it unlikely that there was a systematic difference between W+HHs and HHs according to this measure.

Acoustic Analysis of HHs

Another potential explanation of the results is that HHs given as part of W+HH combinations were acoustically distinct from HHs given as part of HH-only bouts. If this were the case, the different information available to receivers when they heard W + HHs or HHs might have been due to differences in the acoustic structure of the HHs rather than the call combination. To test this hypothesis, we analysed 20 HH-only bouts and 19 W + HH bouts from five different individuals. Each individual contributed between three and six HH bouts from both HH-only bouts and W + HH combinations. We used eight spectral measurements (Supplementary Table S2) to construct two models: a discriminant function analysis (DFA) and a GLMM. The linear DFA with jackknifed prediction correctly classified 69% of HHs as being part of an HH-only bout or a W+HH combination. To test whether these proportions were significant, and to control for individual identity, we created a GLMM in which the eight acoustic measurements served as predictor variables, individual identity and call bout ID were random effects, and call type (HH or W + HH) was the outcome variable. The full model including all eight acoustic measurements as predictor variables did not fit the data significantly better than the null model that included only the random effects ($\chi^2_{8} = 10.22, P = 0.25$). Thus, it appears that HHs given during HH-only bouts did not differ significantly in their acoustic features from HHs given during W + HH combinations.

Call Subtypes

Callers appeared to systematically vary the acoustic structure of whistles depending on whether the call was produced spontaneously or in response. We recorded 56 spontaneous W + HHs and 13 W + HHs given in response that were suitable for acoustic analysis. Many whistles were flat, with a relatively stable frequency over the course of the call (Fig. 3a). Other whistles ('squiggles') showed much greater frequency modulation (Fig. 3b).

Almost all (91%; 51/56) spontaneous W + HHs contained flat whistles. By contrast, 85% (11/13) of W + HHs given in response contained a squiggle whistle. To test the association between whistle type and call order position (spontaneous/response), we performed a GLMM with call whistle type (flat/squiggle) as the dependent measure and call order position as a predictor variable. Call order position was a significant predictor of whistle type ($\beta = 4.0, SE = 0.9, z = 4.5, P = 0.000$). While the sample of observed squiggle whistles was small, all four individuals that produced at least two response W + HHs produced squiggle whistles, suggesting that it is a feature of the bonobo vocal repertoire, not an idiosyncratic vocalization (Fig. 4). Only males were observed to produce squiggle whistles. There did not appear to be an effect of rank on the production of squiggle whistles.

In contrast to the two subtypes of 'whistles', there appeared to be no acoustic differences between spontaneous and response HHs that were likely discernable to listeners. Call unit duration and number of call units in a call bout were the acoustic parameters that best predicted whether a call was given spontaneously or in response. The mean \pm SD duration of spontaneous HH units was 0.192 ± 0.058 s and that for response HH units was 0.154 ± 0.042 s (mean difference = 0.038 s). The mean number of call units in spontaneous HHs was 5.25 ± 4.0 ; the mean number of call units in

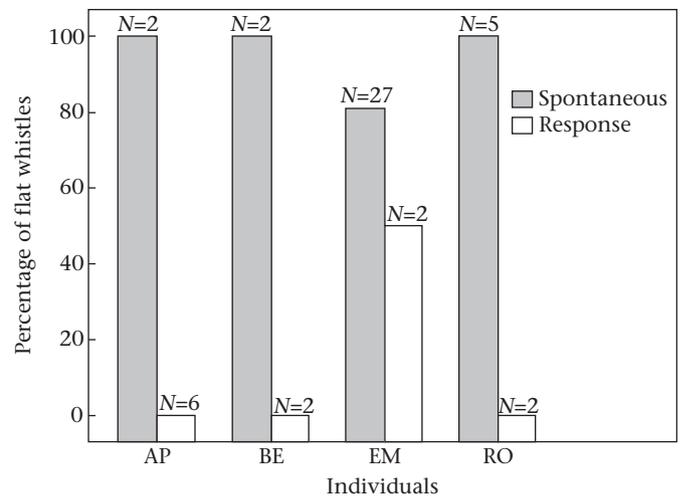


Figure 4. Percentage of flat (versus squiggle) whistles produced as part of spontaneous whistle–high hoot combinations (W + HHs) and response W + HHs. Data are shown for the four individuals who produced at least two spontaneous and two response W + HHs.

response HHs was 4.6 ± 3.3 (mean difference = 0.65 call units). Using call unit duration as the sole predictor, a linear discriminant function (LDF) analysis with jackknifed prediction classified 72 HH sequences as either spontaneous or response calls with 68% accuracy. With number of call units as the sole predictor, calls were classified with 67% accuracy. To test the significance of these predictions we ran a GLMM with spontaneous/response as the dependent measure, call unit duration and number of call units as predictor variables, and ID as a random factor. Unit duration was a significant predictor of the outcome, whereas number of call units was not (call unit duration: $\beta = 16.3, SE = 6.1, z = 2.7, P = 0.007$; number of call units: $\beta = 0.0, SE = 0.1, z = 0.9, P = 0.407$). It is unclear, however, whether such a very small disparity in unit duration (0.038 s) was perceptually significant to bonobos.

In summary, bonobos' use of W + HHs both spontaneously and in response to another caller potentially created an ambiguity for listeners. Callers appeared, however, to systematically vary the acoustic structure of whistles, thereby potentially providing listeners who had just called with the information that the call heard seconds after their vocalization was indeed a response to the call they had just given.

DISCUSSION

Bonobos use call combinations, call exchanges and call subtypes to coordinate their movement between parties. Callers were significantly more likely to travel to a new party after producing W + HHs than after producing HHs alone. They were especially likely to move to another party if their initial vocalizations elicited an answer (i.e. was part of a vocal exchange). Individuals also modified the acoustic structure of their call combinations in a manner that could have allowed listeners to distinguish between those given spontaneously and those given in response to another call. Additionally, callers appeared motivated to receive an answer before travelling to another party, as indicated by their apparent persistence in production of W + HHs.

These communicative abilities may have evolved in bonobos to solve a dilemma confronted by individuals living in fission–fusion societies: how to coordinate movement between individuals in separate parties. Many species use vocalizations to facilitate movement (e.g. Boinski, 1993; Cheney, Seyfarth, & Palombit, 1996),

but the lack of cohesion in fission–fusion societies makes this coordination more difficult (Aureli et al., 2008). In addition, over long distances and with limited visibility, bonobos face a problem common to all social interactions: whenever two individuals approach one another there is uncertainty about the outcome, since the best strategy for each depends on what the other is likely to do (Silk, Kaldor, & Boyd, 2000).

In response to these obstacles, bonobos appear to utilize several features of communication that, taken together, have not previously been documented in nonhuman primates: call combinations that accurately predict the caller's imminent behaviour; possibly persistent call production such that callers tend to produce W + HHs until they receive an apparent response; vocal exchanges in which the first caller's subsequent behaviour is contingent upon the second caller's response; and the selective use of acoustically distinct call subtypes, effectively marking a call combination as an apparent response to another immediately prior vocalization. These four phenomena have been separately observed in other primate species (context-specific call combinations reviewed in Zuberbühler & Lemasson, 2013; persistence: Wich and deVries 2006; Schel, Townsend, Machanda, Zuberbühler, & Slocombe, 2013; Fedurek et al., 2014; exchange-dependent behaviour: Digweed, Fedigan, & Rendall, 2007; acoustic modification in vocal responses: Sugiura, 1998). However, the use of these features in the same communicative event allows bonobos to coordinate movement between foraging parties despite an inherently unpredictable social structure.

Many of the features exhibited by W + HH exchanges are common in birds, both in male–male countersinging and male–female dueting. Nightingales, *Luscinia megarhynchos*, for example, produce multi-element songs in which the presence of one of the elements (the 'trill') signals a caller's motivation to escalate aggression. Furthermore, whether or not the caller receives a response affects its subsequent behaviour (Kunc, Amrhein, & Naguib, 2006; Schmidt, Kunc, Amrhein, & Naguib, 2008; Sprau, Schmidt, Roth, Amrhein, & Naguib, 2010, 2013). Such exchanges are widespread among passerines (e.g. Searcy, Akçay, Nowicki, & Beecher, 2014), and, in many ways, W + HHs resemble these vocal interactions. Call persistence may be an important aspect of W + HH exchanges that is not typically observed in birdsong, but additional research is needed to clarify the existence and underlying cognitive mechanisms of this phenomenon. Nevertheless, the parallels between birdsong and W + HHs clearly demonstrate that the use of call combinations, vocal exchanges and call subtypes is not unique to bonobos, nor are the selective pressures associated with fission–fusion social structure necessary for the evolution of such communicative features.

Bonobos must frequently decide which social partners to associate with and which resources to exploit. These decisions depend on the behaviour, location and motivation of other, out-of-sight individuals, creating an unstable, and potentially confusing, environment. Callers who can signal their imminent behaviour, direct their calls at specific individuals and adjust their behaviour based on the occurrence of call exchanges may decrease the uncertainty associated with fission–fusion dynamics and reduce the costs of group coordination.

Although evidence suggests that the addition of whistles to a series of high hoots signals the caller's motivation to join another party, the function of high hoots produced alone, in the absence of a whistle, remains to be determined. It seems possible that high hoots function to signal the caller's identity and location, and that listeners' responses to both call types depend in part on their relationship to the caller (e.g. baboons: Cheney et al., 1996). Playback experiments have the potential to elucidate these questions. As mentioned previously, whistles are almost always produced in combination with other call types, either with high hoots in the

context of interparty movement or with 'contest hoots' in the context of aggression (Genty et al., 2014). Another future direction would be to compare bouts of contest hoots that contain whistles and those that do not in order to identify features that differ systematically between 'whistle–contest hoot' combinations and 'contest hoots' alone. Such data might allow researchers to assess the impact of 'whistles' in the two different contexts and to draw tentative conclusions about the use of whistles more generally.

Conclusion

Bonobos use call combinations, call exchanges and call subtypes to coordinate movement. Previous research has not documented the integration of these features in nonhuman primate vocalizations. The ability to increase communicative complexity and efficacy by combining existing abilities may have played an important role in the evolution of flexible communication across diverse taxa.

Acknowledgments

We thank the Institut Congolais pour la Conservation de la Nature (ICCN) for permission to conduct research in D.R. Congo. We are grateful to Lys Alcayna for data collection in the field and Robyn Thiessen-Bock for drawing our attention to the W+HH combination. We also want to thank Emily Bray, Andy Gersick and Noah Synder-Mackler for their assistance, feedback and support. Research was funded by a National Science Foundation Graduate Research Fellowship (I.S.) and grants from the Leaky Foundation (I.S.), the Department of Psychology at the University of Pennsylvania (I.S.) and the National Geographic Society (R.M.S.) (Grant number 9115-12).

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.anbehav.2016.10.003>.

References

- Altmann, J. (1974). Observational study of behaviour: Sampling methods. *Behaviour*, 49, 227–267.
- Arnold, K., & Zuberbühler, K. (2006a). The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour*, 72, 643–653. <http://dx.doi.org/10.1016/j.anbehav.2005.11.017>.
- Arnold, K., & Zuberbühler, K. (2006b). Language evolution: Semantic combinations in primate calls. *Nature*, 441(7091). <http://dx.doi.org/10.1038/441303a>, 303–303.
- Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. *Current Biology*, 18(5), R202–R203. <http://dx.doi.org/10.1016/j.cub.2008.01.040>.
- Arnold, K., & Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, 120, 303–309. <http://dx.doi.org/10.1016/j.bandl.2011.10.001>.
- Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., et al. (2008). Fission–fusion dynamics: New research frameworks. *Current Anthropology*, 49, 627–654. <http://dx.doi.org/10.1086/586708>.
- Bermejo, M., & Omedes, A. (1999). Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, 70, 328–357. <http://dx.doi.org/10.1159/000021717>.
- Boinski, S. (1993). Vocal coordination of troop movement among white-faced capuchin monkeys, *Cebus capucinus*. *American Journal of Primatology*, 30, 85–100. <http://dx.doi.org/10.1002/ajp.1350300202>.
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. *Animal Cognition*, 15(3), 327–339. <http://dx.doi.org/10.1007/s10071-011-0456-8>.
- Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biology Letters*, 9(5). <http://dx.doi.org/10.1098/rsbl.2013.0535>.
- Cheney, D. L., Seyfarth, R. M., & Palombit, R. (1996). The function and mechanisms underlying baboon 'contact' barks. *Animal Behaviour*, 52, 507–518. <http://dx.doi.org/10.1006/anbe.1996.0193>.

- Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS One*, 1(1), e73. <http://dx.doi.org/10.1371/journal.pone.0000073>.
- Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, 77, 1387–1396. <http://dx.doi.org/10.1016/j.anbehav.2009.02.016>.
- Clay, Z., & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLoS One*, 6, e18786.
- Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: Syntax before phonology? *Proceedings of the Royal Society B: Biological Sciences*, 281(1788), 20140263. <http://dx.doi.org/10.1098/rspb.2014.0263>.
- Crawley, J. (2014). *Statistics: An introduction using R* (2nd ed.). New York, NY: Wiley.
- Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, 142, 397–421. <http://dx.doi.org/10.1163/1568539054012047>.
- Digweed, S. M., Fedigan, L. M., & Rendall, D. (2007). Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus capucinus*). *American Journal of Primatology*, 69, 829–835. <http://dx.doi.org/10.1002/ajp.20398>.
- Fedurek, P., Donnellan, E., & Slocombe, K. E. (2014). Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behavioral Ecology and Sociobiology*, 68(8), 1345–1355. <http://dx.doi.org/10.1007/s00265-014-1745-4>.
- Genty, E., Clay, Z., Hobaier, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PLoS One*, 9, e84738. <http://dx.doi.org/10.1371/journal.pone.0084738>.
- Gersick, A., Cheney, D. L., Schneider, J., Seyfarth, R. M., & Holekamp, K. (2015). Long-distance communication facilitates cooperation among wild spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, 103, 107–116. <http://dx.doi.org/10.1016/j.anbehav.2015.02.003>.
- Hohmann, G., & Fruth, B. (1994). Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology*, 15, 767–782. <http://dx.doi.org/10.1007/BF02737430>.
- Hohmann, G., & Fruth, B. (2003). Lui Kotal: A new site for field research on bonobos in the Salonga National Park. *Pan African News*, 10, 25–27.
- Kalan, A., & Boesch, C. (2015). Audience effects in chimpanzee food calls and their potential for recruiting others. *Behavioral Ecology and Sociobiology*, 69(10), 1701–1712. <http://dx.doi.org/10.1007/s00265-015-1982-1>.
- Kano, T. (1992). *The last ape: Pygmy chimpanzee behaviour and ecology*. Palo Alto, CA: Stanford University Press.
- Kunc, H., Amrhein, V., & Naguib, M. (2006). Vocal interactions in nightingales, *Luscinia megarhynchos*: More aggressive males have higher pairing success. *Animal Behaviour*, 56, 25–30. <http://dx.doi.org/10.1016/j.anbehav.2005.08.014>.
- Lehmann, J., & Boesch, C. (2004). To fission or to fusion: Effects of community size on wild chimpanzee (*Pan troglodytes verus*) social organisation. *Behavioral Ecology and Sociobiology*, 56, 207–216. <http://dx.doi.org/10.1007/s00265-004-0781>.
- Mitani, J. C., & Nishida, T. (1993). Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour*, 45, 735–746. <http://dx.doi.org/10.1006/anbe.1993.1088>.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 22026–22031. <http://dx.doi.org/10.1073/pnas.0908118106>.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell's monkeys use affixation to alter call meaning. *PLoS One*, 4(11), e7808. <http://dx.doi.org/10.1371/journal.pone.0007808>.
- Ramos-Fernández, G. (2005). Vocal communication in a fission–fusion society: Do spider monkeys stay in touch with close associates? *International Journal of Primatology*, 26, 1077–1092. <http://dx.doi.org/10.1007/s10764-005-6459-z>.
- 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), e76674. <http://dx.doi.org/10.1371/journal.pone.0076674>.
- Schel, A. M., Tranquilli, S., & Zuberbühler, K. (2009). The alarm call system of two species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*). *Journal of Comparative Psychology*, 123(2), 136–150. <http://dx.doi.org/10.1037/a0014280>.
- Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., et al. (2014). Monkey semantics: Two 'dialects' of Campbell's monkey alarm calls. *Linguistics and Philosophy*, 37(6), 439–501. <http://dx.doi.org/10.1007/s10988-014-9155-7>.
- Schlenker, P., Chemla, E., Arnold, K., & Zuberbühler, K. (2016). Pyow-hack revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua*, 171, 1–23.
- Schmidt, R., Kunc, H., Amrhein, V., & Naguib, M. (2008). Aggressive responses to broadband trills are related to subsequent pairing success in nightingales. *Behavioral Ecology*, 19(3), 635–641. <http://dx.doi.org/10.1093/beheco/arn021>.
- Searcy, W. A., Akçay, C., Nowicki, S., & Beecher, M. D. (2014). Aggressive signaling in song. In M. Naguib, L. Barrett, J. Brockmann, S. Healy, J. C. Mitani, T. J. Roper, et al. (Eds.), *Vol. 46. Advances in the study of behaviour* (pp. 89–125). Amsterdam, The Netherlands: Elsevier. <http://dx.doi.org/10.1016/B978-0-12-800286-5.00003-1>.
- Silk, J. B., Kaldor, E., & Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour*, 59, 423432. <http://dx.doi.org/10.1006/anbe.1999.1312>.
- Spehar, S. N., & Di Fiore, A. (2013). Loud calls as a mechanism of social coordination in a fission–fusion taxon, the white-bellied spider monkey (*Ateles belzebuth*). *Behavioral Ecology and Sociobiology*, 67(6), 947–961. <http://dx.doi.org/10.1007/s00265-013-1520-y>.
- Sprau, P., Schmidt, R., Roth, T., Amrhein, V., & Naguib, M. (2010). Effects of rapid broadband trills on responses to song overlapping in nightingales. *Ethology*, 116(4), 300–308. <http://dx.doi.org/10.1111/j.1439-0310.2010.01743.x>.
- Sprau, P., Schmidt, R., Roth, T., Amrhein, V., & Naguib, M. (2013). The predictive value of trill performance in a large repertoire songbird, the nightingale *Luscinia megarhynchos*. *Journal of Avian Biology*, 44(6), 567–574. <http://dx.doi.org/10.1111/j.1600-048X.2013.00113.x>.
- Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, 55, 673–687. <http://dx.doi.org/10.1006/anbe.1997.0602>.
- de Waal, F. B. M. (1988). The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees. *Behaviour*, 106, 183–251.
- White, F. J., Waller, M., Boose, K., Merrill, M. Y., & Wood, K. D. (2015). Function of loud calls in wild bonobos. *Journal of Anthropological Sciences*, 93, 1–13. <http://dx.doi.org/10.4436/JASS.93003>.
- Wich, S. A., & deVries, H. (2006). Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society B: Biological Sciences*, 273(1587), 735–740. <http://dx.doi.org/10.1098/rspb.2005.3320>.
- Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour*, 63, 293–299. <http://dx.doi.org/10.1006/anbe.2001.1914>.
- Zuberbühler, K., & Lemasson, J. (2013). Primate communication: Meaning from strings of calls. In F. Lowenthal, & L. Lefebvre (Eds.), *Language and recursion* (pp. 115–128). Berlin, Germany: Springer.