






Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Chimpanzees combine pant hoots with food calls into larger structures



Maël Leroux ^{a, b, c, *} , Alexandra B. Bosshard ^{a, c} , Bosco Chandia ^b , Andri Manser ^d , Klaus Zuberbühler ^{b, c, e, f} , Simon W. Townsend ^{a, c, g}

^a Department of Comparative Language Science, University of Zürich, Zürich, Switzerland

^b Budongo Conservation Field Station, Masindi, Uganda

^c Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland

^d Department of Evolutionary Biology and Environmental Studies, University of Zürich, Zürich, Switzerland

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

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

^g Department of Psychology, University of Warwick, Warwick, U.K.

ARTICLE INFO

Article history:

Received 10 February 2021

Initial acceptance 15 March 2021

Final acceptance 21 May 2021

MS. number: 21-00099R

Keywords:

call combination

evolution of language

great ape

nonhuman primate

Pan troglodytes schweinfurthii

syntax

A growing body of observational and experimental data in nonhuman primates has highlighted the presence of rudimentary call combinations within the vocal communication system of monkeys. Such evidence suggests the ability to combine meaning-bearing units into larger structures, a key feature of language also known as syntax, could have its origins rooted within the primate lineage. However, the evolutionary progression of this trait remains ambiguous as evidence for similar combinations in great apes, our closest-living relatives, is sparse and incomplete. In this study, we aimed to bridge this gap by analysing the combinatorial properties of the pant hoot–food call combination in our closest-living relative, the chimpanzee, *Pan troglodytes*. To systematically investigate the syntactic-like potential of this structure, we adopted three levels of analysis. First, we applied collocation analyses, methods traditionally used in language sciences, to confirm the combination of pant hoots with food calls was not a random co-occurrence, but instead a consistently produced structure. Second, using acoustic analyses, we confirmed pant hoots and food calls comprising the combination were acoustically indistinguishable from the same calls produced in isolation, indicating the pant hoot–food call combination is composed of individually occurring meaning-bearing units, a key criterion of linguistic syntax. Finally, we investigated the context-specific nature of this structure, demonstrating that the call combination was more likely to be produced when feeding on larger patches and when a high-ranking individual joined the feeding party. Together our results converge to provide support for the systematic combination of calls in chimpanzees. We highlight that playback experiments are vital to robustly disentangle both the function this combination might serve and the similarities with combinations of meaning-bearing units (i.e. syntax) in language.

© 2021 The Author(s). Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Human language is described as the hallmark of our species (Lieberman, 1991). One feature that has been repeatedly highlighted to distinguish language from other nonhuman communication systems is syntax (Bolhuis, Tattersall, Chomsky, & Berwick, 2014; Fitch, 2018). Through combining meaningful units (i.e. morphemes/words) into larger structures with related (compositional) or unrelated (idiomatic e.g. kick the bucket) meanings, human language

harnesses a potentially infinite generative power. While animal communication is undoubtedly far more bounded and unproductive in comparison to language (Hurford, 2011), emerging data in nonhuman animals have recently challenged the assumption that the basic capacity for syntax, i.e. merging meaning-bearing units, is unique to the human lineage (monkeys: Arnold & Zuberbühler, 2006a; Berthet et al., 2019; Coye, Ouattara, Zuberbühler, & Lemasson, 2015; Coye, Zuberbühler, & Lemasson, 2016; birds: Engesser, Ridley, & Townsend, 2016; Suzuki, Wheatcroft, & Griesser, 2016, 2017). For instance, putty-nosed monkeys, *Cercopithecus nictans*, produce two acoustically distinct alarm calls: ‘hacks’ produced

* Corresponding author.

E-mail address: mael.leroux@uzh.ch (M. Leroux).

when encountering aerial threats (e.g. eagle) and 'pyows' produced for a variety of disturbances originating from the ground (Arnold & Zuberbühler, 2006b). Importantly, individuals combine these two calls into larger utterances which function to initiate group travel (Arnold & Zuberbühler, 2006a). Since this behaviour is generally unrelated to antipredator contexts, this combination has been argued to resemble an idiomatic syntactic-like construction (Arnold & Zuberbühler, 2012), in which the meaning of the structure is not directly related to the meaning of its comprising units (akin to the English phrase 'kick the bucket', Hurford, 2011; but see also Schlenker, Chemla, Arnold, and Zuberbühler (2016) for alternative interpretations).

Diana monkeys, *Cercopithecus diana*, have also been shown to combine calls into larger structures. Females produce three distinct social calls: 'L' calls produced in a sociopositive context, 'R' calls in a socionegative context and 'A' calls that differ primarily between individuals (Candiotti, Zuberbühler, & Lemasson, 2012a, 2012b). Notably, the monkeys combine these calls in nonrandom sequences (i.e. 'L-A' and 'R-A') that encode information regarding both the context of production and the identity of the caller (Coye et al., 2016). These results suggest the meaning of the structure is derived from the meaning of the comprising units, comparable to compositional structures in human syntax (i.e. 'duck and cover!', Hurford, 2011; Townsend, Engesser, Stoll, Zuberbühler, & Bickel, 2018).

Such evidence for syntactic-like structures among nonhuman primates, species closely related to humans, has been suggested to indicate the core cognitive abilities facilitating syntax could have emerged as early as in the last common ancestor of monkeys and humans, approximately 45 million years ago (Leroux & Townsend, 2020). However, similar combinatorial phenomena have been demonstrated in more distantly related species, such as birds (Engesser et al., 2016; Suzuki et al., 2016), suggesting these abilities can arise as a product of convergent evolutionary processes. Furthermore, data from great apes, our closest-living relatives, which are key to disentangling shared ancestry from convergence, remain sparse and ambiguous.

One pertinent example comes from gorillas, *Gorilla gorilla*, which have been shown to combine contact calls (e.g. 'A1', 'T2', 'T4') into various sequences when responding to other group members (e.g. 'A1-T4', 'T2-T4'; Hedwig, Mundry, Robbins, & Boesch, 2015; Hedwig, Robbins, Mundry, Hammerschmidt, & Boesch, 2014). The combinations are produced when resting, similarly to 'A1' and 'T2' when produced in isolation, suggesting the meaning of the combinations could be derived from the meaning of their parts (Hedwig et al., 2015). However, there is also a degree of contextual discontinuity as 'T4' calls are produced in isolation in a feeding context (Hedwig et al., 2015). Hence these sequences may represent more read-outs of concomitant shifts in context as opposed to communicating a related but distinct context (as is the case in other species producing combinatorial structures: see Engesser et al., 2016). Similarly, bonobos, *Pan paniscus*, combine different call types, namely 'whistles' and 'high-hoots' into a larger sequence to coordinate movements across groups (Schamberg, Cheney, Clay, Hohmann, & Seyfarth, 2016). Interestingly, receivers were more likely to switch parties after hearing a combination compared to 'high-hoots' alone, suggesting the combination has a distinct function to that of 'high-hoots' in isolation (Schamberg et al., 2016). Although promising, it remains unclear whether 'whistles' occur in isolation and the specific function they might have, rendering it complicated to link the function and meaning of the parts to the combination. Lastly, an extensive overview of call combinations has also been documented in chimpanzees, *Pan troglodytes*, with the identification of at least 88 different call combinations built from the 15 distinct call types making up their vocal repertoire (Crockford & Boesch, 2005). However, this report is highly

descriptive, precluding objective investigation into the properties and functions of the combinations.

While these data are certainly suggestive of a similar capability to systematically combine calls together into larger utterances in great apes, they suffer from the following three limitations. First, in each of the above examples it is not possible to objectively rule out the possibility that the combinations identified are simply independent calls that happen to fall next to each other by chance. Any combinatorial-like properties or features may therefore be illusory. Second, statistical verification that the calls comprising the combination are acoustically indistinguishable from their independently occurring counterparts is lacking. Since the definition of syntax requires the combination of meaning-bearing units together (Suzuki & Zuberbühler, 2019), such an analysis is integral to demonstrating syntactic-like structures in nonhuman animals. Lastly, the accompanying context of production and function of great ape call combinations has seldom been quantified (but see Schamberg, Cheney, Clay, Hohmann, & Seyfarth, 2017). Without this, it becomes problematic to convincingly investigate the semantic relationship between a nonrandom call combination and its individual call building blocks: namely whether the combination is related to the meaning of its comprising parts (compositional structuring) or whether it is distinct from them (idiomatic structuring).

In this study, we aimed to bridge this gap through systematically unpacking the combinatorial potential and function of the pant hoot–food call sequence in wild chimpanzees, *P. t. schweinfurthii*. Preliminary work adopting a high-density data approach identified sequences of pant hoots juxtaposed with food calls as promising candidates for syntactic-like structures (Leroux, 2021). Pant hoots are long-distance calls, produced in a wide variety of contexts (e.g. travelling, resting, feeding) and have been shown to reliably encode the identity of the caller (Fedurek, Zuberbühler, & Dahl, 2016). In contrast, food calls are highly context-specific calls, produced exclusively in relation to feeding (Slocombe & Zuberbühler, 2010) and function, among other things, to communicate the presence of food (Goodall, 1986; Slocombe & Zuberbühler, 2005, 2006). Critically, chimpanzees seem to combine these calls specifically when individuals are joining a feeding party (M. Leroux, personal observation). Given the overlap in contexts in which the individual calls and the combination are given, we reasoned it might represent a compositional-like structure communicating related but additional information regarding the individual and the presence of food.

We adopted three complementary levels of analysis to investigate the combinatorial nature of this structure and its potential meaning and function. To assess whether the pant hoot–food call combination is a simple chance co-occurrence of two calls with each other (i.e. read-outs of context changes) or, rather, a nonrandom combination, we applied collocation analyses to chimpanzee calls, an approach traditionally used in computational linguistics to calculate word co-occurrence probability (Gries, 2013). Then, to confirm that the combination comprises individually occurring meaning-bearing units, we compared the fine acoustic structure of pant hoots and food calls given in isolation with the calls comprising the combination. Lastly, we quantitatively probed the function of the pant hoot–food call sequence by exploring the relative contribution of different socioecological factors on the production of this structure.

METHODS

Ethical Note

The study was approved by the Uganda Wildlife Authority (permit number: COD/96/05), the Uganda National Council for

Science and Technology (permit number: NS47ES) and the Animal Welfare & Ethical Review Body from the University of Warwick, U.K. (permit number: AWERB.36/17–18). The study was purely observational, and all applicable international and national guidelines were followed.

Study Site and Subjects

The study was conducted at the Budongo Conservation Field Station (BCFS), Budongo forest, Uganda. The Sonso community has been habituated to researchers' presence since 1990 and at the time of the study consisted of 72 identifiable individuals, including 43 adults (12 males, 31 females), 11 subadults (two males, nine females), 10 juveniles (eight males, two females) and eight infants (four males, four females; see Appendix Table A1 for the list of individuals; age categories based on Reynolds, 2005).

Data Collection

Audio recordings

The data were collected over two field periods (February–June 2019 and September 2019–March 2020) for a total of 12 months. We followed 31 adults using focal animal sampling for a duration of 2 h, totalling 198 focal hours. During a focal follow, all vocalizations were recorded using a Marantz PMD661 mk3 handheld digital audio recorder (Marantz, Kanagawa, Japan; sample rate 44.1 kHz, resolution 32 bits, .wav format) connected to a directional Sennheiser ME66/K6 microphone (Sennheiser, Wedemark, Germany). For each call recorded, the caller ID and context of production were noted after the calls. Vocalizations were extracted using Adobe Audition (2015.0) and, in line with previous work in great apes, a maximum time interval of ≤ 2 s between two calls produced by the same caller was used to define a call combination (Crockford & Boesch, 2005; Hedwig et al., 2014).

Feeding events

A pilot study highlighted that the production of pant hoot–food call combination occurred overwhelmingly during feeding (Leroux, 2021); hence, we restricted our focus to this specific context. We collected data on 145 feeding events. For each event, all occurrence sampling was used to collect pant hoot–food call combinations produced by individuals within the feeding party (defined as all individuals within a 30 m radius of the focal subject, Slocombe et al., 2010). Owing to the high degree of chorusing during these feeding events as well as low visibility in some tree species, we excluded call combinations that were not of sufficient quality to confidently verify the caller's ID and acoustically validate the ≤ 2 s silence interval criterion between the calls constituting the combination (129/187 combinations removed). In 13 of the 145 feeding events, none of the callers could be identified, leading to the exclusion of these 13 events for further analysis. For the remaining 132 events, we noted the type of food consumed (ripe fruits, unripe fruits, leaves or flowers) and the duration of the feeding event as a proxy of the patch size (i.e. chimp minutes) in line with previous work investigating the dynamics of feeding behaviour in chimpanzees (Fedurek & Slocombe, 2013; Slocombe et al., 2010; White & Wrangham, 1988). On some occasions we likely underestimated the total feeding time as parties were often already present when a focal individual arrived at a feeding patch or stayed after it left. Finally, we noted whether any individuals joined the party during the feeding event. Only feeding events in which we could document all the above-mentioned variables were considered for further statistical analysis ($N_{\text{Feeding events}}=98$).

Dominance Hierarchy Among Group Members

Dominance hierarchy was calculated for adult males based on the production of pant grunt vocalizations, which are known to signal subordination from the caller to the receiver and are therefore a reliable indicator of dominance relationships in chimpanzees (Goodall, 1986). Pant grunt data (December 2017–March 2020) were extracted from the long-term data collected by field assistants at BCFS. Ranks were calculated using Elo-rating procedures (Neumann, Duboscq, Dubuc, & Ginting, 2011; see Appendix Table A2 for the Elo-ratings).

Collocation Analysis

To investigate the nonrandom nature of the combination of pant hoots with food calls (i.e. not concatenated purely by chance, or simple read-outs of changes in contexts, akin to a running commentary of events), we compared the rate of occurrence of the pant hoot–food call combination with other two-call combinations (bigrams) using focal data. Specifically, we calculated call dependencies within 75 distinct bigrams using a multiple distinctive collocation analysis in R (version 3.6.0, R Core Team, 2019; Gries 2007). Collocation analysis is a constructional-based technique commonly used in language sciences which compares co-occurrences of specific words (here call types) with each other (Church, Gale, Hanks, & Hindle, 1991; Gries & Stefanowitsch, 2004; Kennedy, 1991; Nesselhauf, 2005; see Bosshard et al., 2021 for its application to animal data sets). Collocation analysis aims to identify how exclusively two units, i.e. words, associate with each other, simply put, the relative attraction between words. For example, if A and B combine (A-B), collocation analysis compares the frequency of A-B with the frequencies of all other bigrams A and B independently appear in (e.g. A-C, B-C, A-D, B-D ...). A multiple distinctive collocation analysis tests this attraction between words using one-tailed exact binomial tests on each possible bigram combination (Gries, 2014) and the log-transformed results provide an estimate of the degree of exclusivity or relative attraction of words (here call types) with one another, i.e. positive if the observed value is higher than expected and negative if the bigram happens less than expected by chance.

However, one specific feature of animal vocal data sets is their relatively smaller size in comparison to human corpora, leading to low frequencies of occurrence of certain bigrams. To control for this, and to further strengthen the robustness of the analysis, we ran a mutual information collocation analysis on the pant hoot–food call combination which gives a strong weight to low-frequency bigrams (Church & Hanks, 1990).

Acoustic Analysis

To verify the acoustic similarities between food calls and pant hoots given in isolation and those in combinations, we acoustically analysed calls from six adult individuals (four males, two females). Specifically, we selected 11 call combinations, 14 pant hoots in isolation and 34 food calls that were suitable for acoustic analysis. Given food calls almost always occur in short bouts and therefore are not technically independent from one another, and to avoid biasing the analysis, we only selected one tonal element per food call bout with a good signal-to-noise ratio. Pant hoots are long-distance contact calls composed of four different phases: introduction, build-up, climax and let-down, each of which can also potentially comprise multiple elements (Slocombe & Zuberbühler, 2010; see Fig. 1). Since pant hoots' acoustic structure has been shown to differ according to its context of production (Fedurek et al., 2016), we only considered pant hoots produced during

feeding. Finally, to avoid similar pseudoreplication issues associated with selecting more than one element per bout and since build-up and let-down phases were not always clearly identifiable, we conducted acoustic analyses only on the first clear tonal element of introduction and climax phases (Fig. 1).

A total of 23 spectral and temporal acoustic parameters were extracted from each element using a custom-built script for PRAAT (version 6.0.33; see Reby & McComb, 2003; Watson, Townsend, & Range, 2018). To validate the accuracy of pitch tracking we compared the time-varying numerical representation of the F0 contour with the F0 from the corresponding spectrogram (Charlton, Zhihe, & Snyder, 2009).

For each element separately, we checked for multicollinearity of acoustic parameters and excluded variables with the greatest variance inflation factor ($VIF > 5$) seriatim until a set of six uncorrelated parameters was obtained (i.e. no greater than the number of individuals tested to avoid overfitting the model, see Table 1; all VIFs in final set were < 2). For parameters that were not normally distributed and could be improved, we applied log transformations.

To test the acoustic similarities between calls produced in isolation and in combination, we used permuted discriminant function analyses (pDFA) in R (Mundry & Sommer, 2007). This variant of a classic DFA is advantageous for two reasons: it allows the use of multiple data points per individual and it can handle unbalanced data sets as is the case here (see Watson et al., 2018 for more details). We conducted one pDFA for each of the three call elements considered (see Table 1).

Socioecological Drivers of Call Combination Production

To explore the relative influence of social and ecological variables on the production of the pant hoot–food call combination, we performed a generalized linear mixed model (GLMM) with a binomial error structure and a logit link function using the glmer function, lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). We modelled the production of pant hoot–food call combinations (0/1) as a binary response variable. As ecological predictors, we fitted food quality and feeding event duration (min). As some food types were infrequently consumed, we coded food quality as a binary fixed effect: low (leaves/flowers) versus high (fruits) quality with quality based on feeding time budget

calculated for chimpanzees of the Budongo forest (Tweheyo, Lye, & Weladji, 2004). As social predictors, we fitted whether a joining event occurred or not (0/1). Since food calls in chimpanzees have been demonstrated to be directed at specific individuals, especially high-ranking males (Schel, Machanda, Townsend, Zuberbühler, & Slocombe, 2013), we specified the presence (0/1) of high-ranking individuals (alpha and beta males) within joining events as a nested fixed effect. Finally, as several feeding events could occur on one day, we fitted date as a random effect.

We checked model assumptions using the DHARMA package in R (Hartig, 2020). None of the independent variables were found to be collinear ($VIFs < 2.5$). The model was not overdispersed ($P = 0.81$), no outliers were detected ($P = 1$) and visual inspection of the Q–Q plots confirmed the normality of the residuals (Kolmogorov–Smirnov test: $P = 0.82$).

RESULTS

Overview

We collected a total of 187 pant hoot–food call combinations from 33 of 42 adult subjects exclusively in feeding contexts. Of the nine individuals who did not produce the combination, all were peripheral females except for the alpha male. We also only recorded the beta male producing the combination once. This low frequency of pant hoot–food call combinations in the alpha and beta males seems to be specific to the sequence itself given they produced the independent comprising calls (food calls and pant hoots) at similar frequencies as other individuals (see Appendix Table A3 for an overview of the independent call production frequencies across males).

Collocation Analyses

We implemented collocation analyses to verify the combination was not simply a chance co-occurrence of calls. Of the 75 bigrams collected during the study period, 23 (31%) were pant hoot–food call combinations. Multiple distinctive collocation analysis demonstrated a significant relative attraction between pant hoots and food calls with the highest relative attraction value among all possible bigrams ($pbin_ph_fc = 9.49$, $P < 0.001$; see Appendix

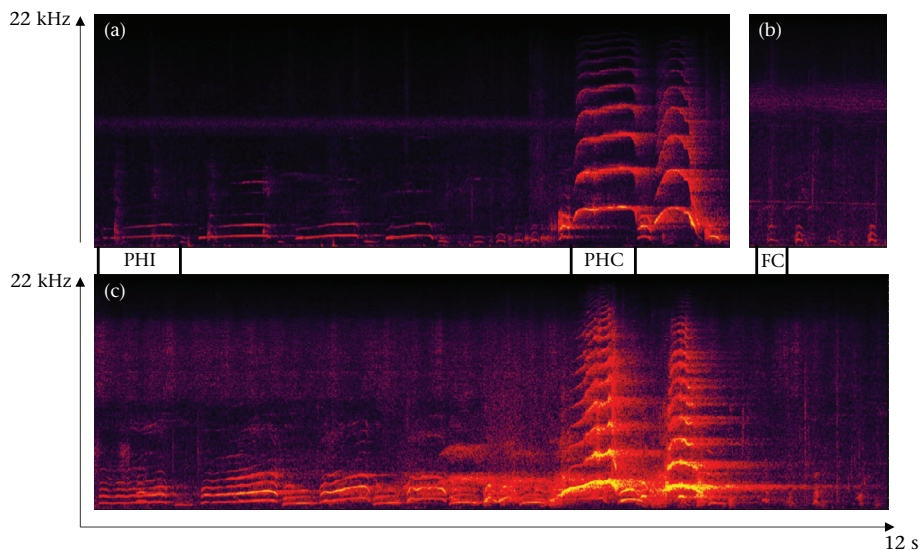


Figure 1. Spectrograms of (a) a pant hoot and (b) a food call produced in isolation and (c) a pant hoot–food call combination. Acoustic analyses were performed on the first element of introduction (PHI) and climax (PHC) of the pant hoot and the first tonal element of a food call bout (FC) for both the calls in isolation and in combination.

Table 1
Acoustic parameters extracted and used for the pDFA analyses of pant hoot introduction (PHI), pant hoot climax (PHC) and food calls (FC)

Vocal parameters	Definition	PHI	PHC	FC
Duration	Duration of the element	x	x	x
F0 start	The value of F0 at the start of the element	x		x
F0 end	The value of F0 at the end of the element		x	
% Time max F0	The percentage of the total duration for which F0 was at maximum	x	x	x
FM rate	The number of complete cycles (peak-to-trough-to-peak) of F0 modulation/s		x	
Q25%	The frequency values at the upper limit of the first quartiles of energy, measured on a linear amplitude spectrum applied to the entire element			x
% EfPeak	The percentage of the total element duration where energy value of the frequency with the highest power/energy of the element was maximum			x
% Time max intensity	The percentage of the total element duration when the intensity was maximum	x	x	x
AM var	The mean variation/s of the intensity contour of the element, calculated as the cumulative variation in amplitude divided by the element duration	x	x	
AM rate	The number of complete cycles of amplitude modulation/s of intensity contour of the element	x		

Table A4). The reversed order combination occurred only twice and the relative attraction between the two calls in the reversed order was not significant ($p_{bin_fc-ph} = 1.13$, $P > 0.05$; see Appendix **Table A4**).

Owing to our small sample size and the rare occurrence of certain bigrams in the data set, we ran a mutual information collocation analysis to increase the model's sensitivity to less frequent sequences. When controlling for this, however, we still found a significant relative attraction between pant hoots and food calls (collocational strength = 1.37, $P < 0.05$).

Acoustic Analysis

The pDFA indicated that introduction and climax elements from pant hoots in isolation and in combination were acoustically identical (pDFA_A: $N_{individuals} = 6$, $N_{calls} = 25$; correctly cross-classified: 39.3%; expected correctly cross-classified: 48.06%; $P = 0.69$; pDFA_B: $N_{individuals} = 6$, $N_{calls} = 23$; correctly cross-classified: 38.36%; expected correctly cross-classified: 47.6%; $P = 0.69$; **Fig. 2a** and **b**). Similarly, food calls in isolation and in combination could not be discriminated based on their acoustic structure (pDFA_C: $N_{individuals} = 6$, $N_{calls} = 45$; correctly cross-classified: 56.9%; expected correctly cross-classified: 50.3%; $P = 0.27$; **Fig. 2c**).

Socioecological Drivers of Call Combination Production

Of 58 combinations for which both the caller ID and the ≤ 2 s time interval criterion could be confidently verified, 40 (69%) were produced when an individual joined a feeding party and 15 (26%) when an individual arrived at an unexploited feeding patch. In the three remaining cases (5%), the combination was produced while feeding, after another party vocalized nearby.

Overall, both ecological and social variables influenced the production of call combinations. A GLMM with a binomial error structure indicated that the production of pant hoot–food call combinations increased with the duration of the feeding event ($P = 0.009$; see **Table 2**, **Fig. 3a**). In addition, the likelihood of producing call combinations was positively associated with the joining of individuals, specifically when high-ranking individuals were part of the joining party ($P = 0.018$; see **Table 2**, **Fig. 3b**; see also Appendix **Table A5**).

DISCUSSION

Using a combination of behavioural, acoustic and computational approaches we have shown that chimpanzees concatenate two independently occurring calls (pant hoots and food calls) in nonrandom ways in highly specific behavioural contexts. Critically,

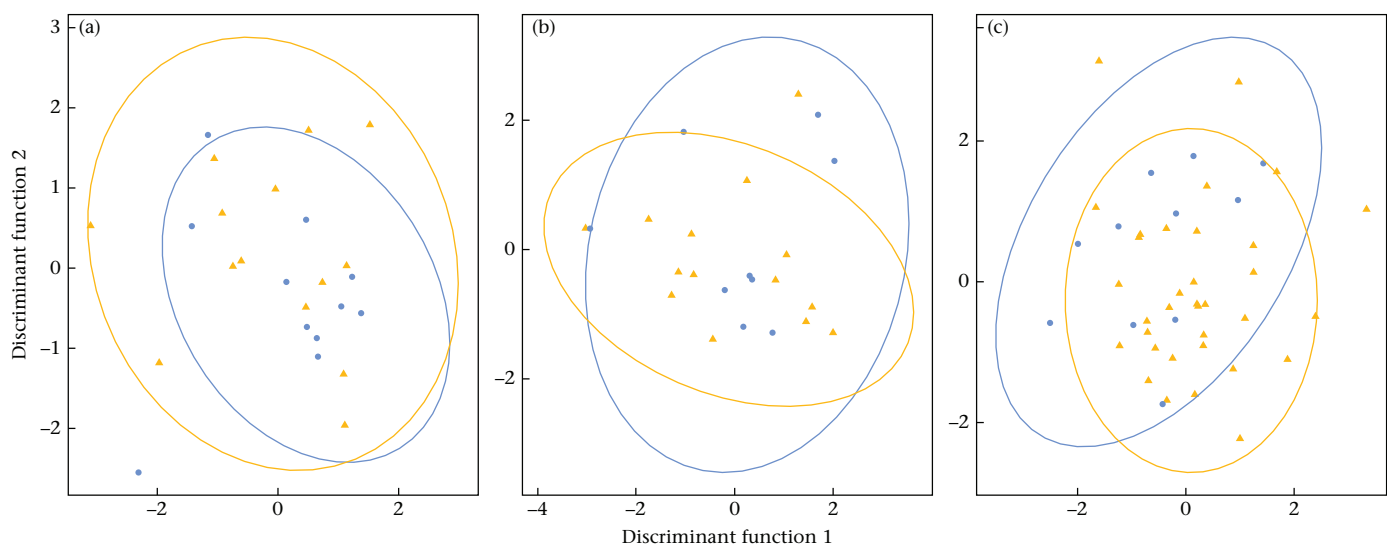


Figure 2. Territorial plot of the first two discriminant functions for (a) pant hoot introduction, (b) pant hoot climax and (c) food call elements occurring in combination (blue dots) or in isolation (yellow triangles). Ellipses represent clusters assuming a multivariate normal distribution with a 95% confidence level.

Table 2

GLMM indicating the influence of joining (0/1) including whether the individual was high ranking or not, duration of feeding event (min) and food quality (high/low)

	Estimate	SE	z	P
(Intercept)	-2.754			
Joining	1.026	0.646	1.587	0.11
Duration	0.022	0.008	2.623	0.009
Food quality	-1.560	0.850	-1.834	0.067
High ranking within joining	2.004	0.845	2.369	0.018

Significant *P* values are shown in bold.

acoustic analyses demonstrated that calls comprising the combination did not differ statistically from the same calls produced in isolation. This is particularly relevant as it suggests that rather than being a distinct, novel call type, the sequence is a product of already existing call units that have subsequently been recombined. Furthermore, by borrowing methods from computational linguistics that have been traditionally used to identify the relative strength of attraction and repulsion of pairs of words (Gries, 2013), we have demonstrated that this sequencing of calls is unlikely to be a random co-occurrence but instead pant hoots and food calls are strongly attracted to one another, more than expected by chance. An alternative interpretation could be that our collocation results are driven more simply by an attraction of single calls to the same context. However, we are confident we can rule this out since, if attraction to the same context was the driving force promoting the concatenation of two calls, we would not expect combinations to be stereotyped in their ordering, as is the case with pant hoot–food call sequences: food calls overwhelmingly follow pant hoots. Furthermore, it could also be that frequent co-occurrences of calls merely result from each call in the combination occurring frequently in isolation; if pant hoots are common and food calls are common, pant hoot–food call combination occurrence is additionally inflated. This alternative mechanism is also unlikely since there would again be no a priori reason to expect calls to be structured in a particular order, which is the case with the pant hoot–food call sequence. Moreover, high-ranking males rarely, if ever, produced the pant hoot–food call combination but did

produce the individual call components at a rate comparable with other males. As a consequence, we are confident that the collocation analyses we implemented identified nonrandom, communicatively relevant, combinations. To date, the identification of combinatorial structures in nonhuman animals has relied on subjective and qualitative methods which cannot rule out more parsimonious explanations, such as loose adjacency, where calls just happen to fall next to one another (e.g. Candiotti, Zuberbühler, & Lemasson, 2012a; Collier, Townsend, & Manser, 2017; Crockford & Boesch, 2005; Hedwig et al., 2015; Ouattara, Lemasson, & Zuberbühler, 2009). By quantifying the frequency of occurrence of two-call sequences throughout the chimpanzee repertoire, we have provided a more objective and systematic assessment of combinatorics. When applying this approach more broadly, it ultimately becomes feasible to compare the combinatorial dynamics of different groups, populations and even species, data that are central to a more detailed understanding of the diversity and evolution of combinatoriality.

Lastly, contextual data indicate the combination is produced under specific behavioural conditions. First, we found a positive effect of time spent feeding on call combination production, which suggests chimpanzees are more likely to produce these structures when feeding on larger food patches (Slocombe et al., 2010). Second, we also found that call combination production was reliably associated with the joining of individuals at a feeding party, suggesting it potentially serves a more social, feeding coordination, function. Of particular interest is the fact that this effect is overwhelmingly driven by high-ranking individuals joining. Exactly why this is the case remains speculative, but insights may be garnered by considering the respective information communicated by the individual call units comprising the pant hoot–food call structure.

Pant hoot vocalizations are long-distance calls produced in isolation in a variety of contexts including feeding or travelling and have been shown to reliably encode caller ID (Fedurek et al., 2016; Kojima, Izumi, & Ceugniet, 2003; Mitani, Gros-Louis, & Macedonia, 1996). In contrast, food calls are highly context specific, produced exclusively during feeding and used to communicate the presence of food (Goodall, 1986; Slocombe & Zuberbühler, 2005, 2006).

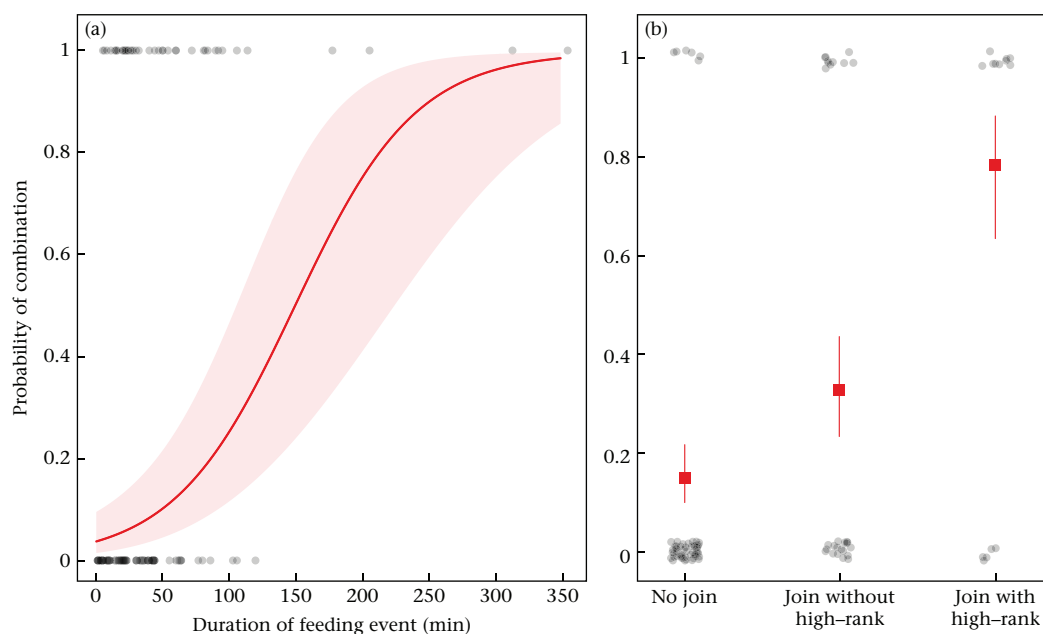


Figure 3. Production of pant hoot–food call combinations (0/1) according to (a) the duration of the feeding event (min) and (b) the joining of high-ranking individuals (0/1). Line (a) and points (b) depict model predictions. Shaded area (a) and bars (b) illustrate the standard errors. Grey dots show the raw data.

Furthermore, previous research has shown food calls in chimpanzees are reliably associated with the presence of valuable social partners (e.g. closely bonded and high-ranking individuals) and are more likely to be produced when feeding on large patches (Schel et al., 2013; Slocombe et al., 2010). By combining the two calls together, chimpanzees can therefore communicate information regarding both the identity of the caller and the context of production, i.e. the presence and characteristics of a feeding patch, simultaneously. This could be beneficial in allowing individuals to potentially attract coalition partners to feeding patches who might otherwise not have visual access given the dense nature of the chimpanzee habitat and the fission–fusion dynamics of the chimpanzee social system (Goodall, 1986; Hobaiter, Byrne, & Zuberbühler, 2017). Unfortunately, with our data set we were not able to explore, at a fine-grained level, how social relationships with joining individuals influenced call combination production. However, we predict a positive effect of the call combination on the joining of preferred social partners, akin to the effect highlighted in food calls (Slocombe et al., 2010). In addition to attracting social partners such combinations could serve an appeasement function, informing high-ranking individuals, who tend to monopolize food resources (Wrangham, 1977), of the presence of food and who has encountered it. For example, in macaques, *Macaca mulatta*, individuals failing to call upon the discovery of a food source received more aggression from other group members or, in the more extreme case of females, led to reduced access to food (Hauser & Marler, 1993). Furthermore, it has been shown that food calling in chimpanzees, in addition to advertising the presence of food, may help to diffuse aggression (Ischer, Zuberbühler, & Fedurek, 2020). Hence, by adding a call with a rich identity signature, the pant hoot–food call combination could reduce the signaller's risk of future aggression. This potential appeasement function is somewhat supported by the fact that call combination production seemed to be overwhelmingly influenced by the joining of high-ranking individuals (either when the caller was already feeding on the tree when the high-ranking individual joined or the caller joined in parallel with the high-ranking individual). This appeasement function hypothesis could be tested by comparing the behaviour of high-ranking individuals joining when the call combination is produced with situations in which it is not. The prediction here would be that high-ranking males are more likely to exhibit aggressive behaviours when the call combination is not produced versus instances when it is produced. Ultimately, while more work on this specific combination is needed to verify its potential function, these results further develop our understanding of the complexity of chimpanzee communication. Combining calls allows individuals to transmit more specific information and ultimately better negotiate their dynamic social worlds.

Aside from shedding light on the potential function of the combination, the overlap in contexts between the pant hoots and the food calls produced in isolation and in combination is also suggestive of a certain semantic transparency in the combination, i.e. the meaning of the whole seems to be related to the meaning of the parts. Moreover, this, in conjunction with the collocation analyses suggesting the combination is a nonrandom structure and acoustic data indicating the comprising units are acoustically identical to individually occurring meaning-bearing counterparts, highlights this combination as a candidate example of a combinatorial construction with syntactic-like properties (see Leroux & Townsend, 2020).

Comparable constructions comprising meaning-bearing units have been documented in other nonhuman primates and non-primate mammals. As previously described, female Diana monkeys, for example, combine contact calls encoding the identity of the signaller with two different calls encoding the context of

production (Coye et al., 2016), while in banded mongooses, *Mungos mungo*, feeding individuals combine an initial noisy segment encoding ID with a temporally varying harmonic unit encoding the precise context of production, specifically whether individuals are moving or searching for food (Jansen, Cant, & Manser, 2012). These structures in banded mongoose have even been compared to simple subject–predicate compositions in human language, a specific form of syntactic structure with the first call acting as referential to ID (similarly to names in language, i.e. Jane) and the harmonic units acting as arguments referencing the context (i.e. digging/moving, Collier, Bickel, van Schaik, Manser, & Townsend, 2014). To our knowledge, our study provides intriguing evidence for such ID–context constructions in a great ape vocal system and hence represents a key data point central to reconstructing the evolutionary roots of linguistic syntax, a hallmark of language.

While promising, playback experiments are needed to shed additional light on both the function of this structure and its subsequent relationship to similar structures in language. In line with previous work, it would be important to contrast responses to the combination with both the individually occurring calls and artificial combinations to probe not only how receivers are potentially binding the meaning of the comprising parts but also to rule out more parsimonious explanations such as priming (one call intensifies the response to another call) or stimulus intensity effects (two calls are more arousing than one, Coye et al., 2015, 2016; Engesser et al., 2016; Suzuki et al., 2016). Finally, playing back the reversed order of the call combination in addition to novel structures would help further disentangle how information encoded in the combination is extracted by listeners (Suzuki et al., 2016, 2017).

Conclusion

Our results providing promising evidence for combinatorial structuring in wild chimpanzees further expand the known phylogenetic distribution of such abilities to include our closest-living relatives. This is critical as the paucity of such evidence in great apes has complicated the reconstruction of the evolutionary progression of syntax in humans (Leroux & Townsend, 2020). The data presented here suggest the origins of syntax might be traceable at least as far back as the last common ancestor of humans and chimpanzees and perhaps even further. However, in addition to experimental validation of the meaning of call combinations, such a conclusion also reasonably hinges on comparable data in other great apes, specifically bonobos, the other species from the *Pan* genus, who are as closely related to humans as chimpanzees. While existing data point towards similar combinatorial propensities (Clay & Zuberbühler, 2011; Schamberg et al., 2016, 2017), a more objective quantification of bonobo combinatorics is needed to rule out these structures are not simple chance occurrences. We hope future research in animal vocal sequences will apply the systematic method used in this study to unambiguously identify combinatorial structures in nonhuman animals.

Author Contributions

Maël Leroux: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing; Alexandra B. Bosshard: Formal analysis, Writing – review & editing; Bosco Chandia: Data curation, Writing – review & editing; Andri Manser: Formal analysis, Writing – review & editing; Klaus Zuberbühler: Resources, Writing – review & editing; Simon W. Townsend: Conceptualization, Methodology, Formal analysis, Resources, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgments

We thank UWA, UNCST and the President's office for permission to conduct the study. We are grateful for the constant support provided by all BCFS staff and the Royal Zoological Society of Scotland (RZSS). We thank Adrian Soldati and two anonymous referees for comments on the manuscript and Guanghao You, Chundra Cathcart, Balthasar Bickel and Sabine Stoll for useful discussions on collocation analyses. This work was supported by the Swiss National Science Foundation (PP00P3_163850) to S.W.T. and the NCCR Evolving Language (Swiss National Science Foundation Agreement #51NF40_180888). We declare no conflict of interest.

References

- Arnold, K., & Zuberbühler, K. (2006a). Semantic combinations in primate calls. *Nature*, 441, 303. <https://doi.org/10.1038/441303a>
- Arnold, K., & Zuberbühler, K. (2006b). The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour*, 72, 643–653. <https://doi.org/10.1016/j.anbehav.2005.11.017>
- Arnold, K., & Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, 120(3), 303–309. <https://doi.org/10.1016/j.bandl.2011.10.001>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme 4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Berthet, M., Mesbahi, G., Pajot, A., César, C., Neumann, C., & Zuberbühler, K. (2019). Titi monkeys combine alarm calls to create probabilistic meaning. *Science Advances*, 5, Article eaav3991.
- Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How could language have evolved? *PLOS Biology*, 12(8), Article e1001934. <https://doi.org/10.1371/journal.pbio.1001934>
- Bosshard, A. B., Leroux, M., Lester, N. A., Bickel, B., Stoll, S., & Townsend, S. W. (2021). From collocations to call-ocations: Using linguistic methods to quantify animal call combinations. *BioRxiv*. <https://doi.org/10.1101/2021.06.16.448679>
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012a). Context-related call combinations in female Diana monkeys. *Animal Cognition*, 15(3), 327–339. <https://doi.org/10.1007/s10071-011-0456-8>
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012b). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, 8(3), 382–385. <https://doi.org/10.1098/rsbl.2011.1182>
- Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009). Vocal cues to identity and relatedness in giant pandas (*Ailuropoda melanoleuca*). *Journal of the Acoustical Society of America*, 126(5), 2721–2732. <https://doi.org/10.1121/1.3224720>
- Church, K., Gale, W., Hanks, P., & Hindle, D. (1991). *Using statistics in lexical analysis*. In U. Zernik (Ed.), *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon* (pp. 115–164). Lawrence Erlbaum Associates.
- Church, K., & Hanks, P. (1990). Word association norms, mutual information and lexicography. *Computational Linguistics*, 16(1), 22–29.
- Clay, Z., & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLOS One*, 6(4), Article e18786. <https://doi.org/10.1371/journal.pone.0018786>
- 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, 20140263. <https://doi.org/10.1098/rspb.2014.0263>
- Collier, K., Townsend, S. W., & Manser, M. B. (2017). Call concatenation in wild meerkats. *Animal Behaviour*, 134, 257–269. <https://doi.org/10.1016/j.anbehav.2016.12.014>
- Coye, C., Ouattara, K., Zuberbühler, K., & Lemasson, A. (2015). Suffixation influences receivers' behaviour in non-human primates. *Proceedings of the Royal Society B: Biological Sciences*, 282(1807), 20150265. <https://doi.org/10.1098/rspb.2015.0265>
- Coye, C., Zuberbühler, K., & Lemasson, A. (2016). Morphologically structured vocalizations in female Diana monkeys. *Animal Behaviour*, 115, 97–105. <https://doi.org/10.1016/j.anbehav.2016.03.010>
- Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, 142, 397–421.
- Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences*, 113(21), 5976–5981. <https://doi.org/10.1073/pnas.1600970113>
- Fedurek, P., & Slocombe, K. E. (2013). The social function of food-associated calls in male chimpanzees. *American Journal of Primatology*, 75(7), 726–739. <https://doi.org/10.1002/ajp.22122>
- Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016). Sequential information in a great ape utterance. *Scientific Reports*, 6(December). <https://doi.org/10.1038/srep38226>
- Fitch, T. W. (2018). Bio-Linguistics: Monkeys break through the syntax barrier. *Current Biology*, 28(12), R695–R697. <https://doi.org/10.1016/j.cub.2018.04.087>
- Goodall, J. (1986). *The chimpanzees of Gombe: patterns of behavior*. Cambridge, MA: Harvard University Press.
- Gries, S. T. (2013). 50-Something years of work on collocations. *International Journal of Corpus Linguistics*, 18(1), 137–166. <https://doi.org/10.1075/ijcl.18.1.09gri>
- Gries, S. T. (2014). Collostructional analysis: Computing the degree of association between words and words/constructions. <http://stgries.faculty.linguistics.ucsb.edu/Teaching/Groningen/Coll.Analysis.R>
- Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis. *International Journal of Corpus Linguistics*, 9(1), 97–129. <https://doi.org/10.1075/ijcl.9.1.06gri>
- Hartig, F. (2020). DHARMA: Residual diagnostics for hierarchical (multi-level/mixed) regression models. *R Package*. <https://cran.r-project.org/web/packages/DHARMA/vignettes/DHARMA.html>
- Hauser, M. D., & Marler, P. (1993). Food-associated calls in rhesus macaques (*Macaca mulatta*): II. Costs and benefits of call production and suppression. *Behavioral Ecology*, 4(3), 206–212. <https://doi.org/10.1093/beheco/4.3.206>
- Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015). Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition*, 18(2), 423–435. <https://doi.org/10.1007/s10071-014-0812-6>
- Hedwig, D., Robbins, M. M., Mundry, R., Hammerschmidt, K., & Boesch, C. (2014). Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour*, 151(8), 1091–1120. <https://doi.org/10.1163/1568539X-00003175>
- Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, 71(96), 1–13. <https://doi.org/10.1007/s00265-017-2325-1>
- Hurford, J. (2011). *The origins of grammar: Language in the light of evolution II*. Oxford: Oxford University Press.
- Ischer, G., Zuberbühler, K., & Fedurek, P. (2020). The relationship between food calling and agonistic behaviour in wild chimpanzees. *Behavioural Processes*, 178(June), 104182. <https://doi.org/10.1016/j.beproc.2020.104182>
- Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. *BMC Biology*, 10, 97. <https://doi.org/10.1186/1741-7007-10-97>
- Kennedy, G. (1991). Between and through: The company they keep and the functions they serve. In K. Aijmer, & B. Altenberg (Eds.), *English Corpus Linguistics* (pp. 95–110). New York, NY: Longman.
- Kojima, S., Izumi, A., & Ceugniet, M. (2003). Identification of vocalizers by pant hoots, pant grunts and screams in a chimpanzee. *Primates*, 44(3), 225–230. <https://doi.org/10.1007/s10329-002-0014-8>
- Leroux, M. (2021). Call combinations in chimpanzees. Unpublished raw data.
- Leroux, M., & Townsend, S. W. (2020). Call combinations in great apes and the evolution of syntax. *Animal Behavior and Cognition*, 7(2), 131–139. <https://doi.org/10.26451/abc.07.02.07.2020>
- Lieberman, P. (1991). Uniquely human: the evolution of speech, thought, and selfless behavior. Cambridge, MA: Harvard University Press.
- Mitani, J. C., Gros-Louis, J., & Macedonia, J. M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. *International Journal of Primatology*, 17(4), 569–583.
- Mundry, R., & Sommer, C. (2007). Discriminant function analysis with non-independent data: Consequences and an alternative. *Animal Behaviour*, 74(4), 965–976. <https://doi.org/10.1016/j.anbehav.2006.12.028>
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam, Netherlands: John Benjamins.
- Neumann, C., Dubocsq, J., Dubuc, C., & Ginting, A. (2011). Assessing dominance hierarchies: Validation and advantages of progressive evaluation with elo-rating. *Animal Behaviour*, 82(4), 911–921. <https://doi.org/10.1016/j.anbehav.2011.07.016>
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, 106(51), 22026–22031. Retrieved from <http://www.pnas.org/content/106/51/22026.full.pdf>
- Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, 65(3), 519–530. <https://doi.org/10.1006/ajbe.2003.2078>
- Reynolds, V. (2005). *The chimpanzees of the Budongo forest: ecology, behaviour and conservation*. Oxford, U.K.: Oxford University Press.
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, 122, 109–116. <https://doi.org/10.1016/j.anbehav.2016.10.003>
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2017). Bonobos use call combinations to facilitate inter-party travel recruitment. *Behavioral Ecology and Sociobiology*, 71(4). <https://doi.org/10.1007/s00265-017-2301-9>
- Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee food calls are directed at specific individuals. *Animal Behaviour*, 86(5), 955–965. <https://doi.org/10.1016/j.anbehav.2013.08.013>
- Schlenker, P., Chemla, E., Arnold, K., & Zuberbühler, K. (2016). Pyow-hack revisited: Two analyses of Putty-nosed monkey alarm calls. *Lingua*, 171, 1–23. <https://doi.org/10.1016/j.lingua.2015.10.002>
- Slocombe, K. E., Kaller, T., Turman, L., Townsend, S. W., Papworth, S., Squibbs, P., et al. (2010). Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behavioral Ecology and Sociobiology*, 64(12), 1959–1966.
- Slocombe, K. E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, 15(19), 1779–1784. <https://doi.org/10.1016/j.cub.2005.08.068>
- Slocombe, K. E., & Zuberbühler, K. (2006). Food-associated calls in chimpanzees: Responses to food types or relative food value? *Animal Behaviour*, 72(5), 989–999.

- Slocombe, K. E., & Zuberbühler, K. (2010). Vocal communication in chimpanzees. In E. V. Lonsdorf, S. R. Ross, & T. Matsuzawa (Eds.), *The Mind of the Chimpanzee* (pp. 192–207). Chicago, IL: The University of Chicago Press.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7, 10986. <https://doi.org/10.1038/ncomms10986>
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2017). Wild birds use an ordering rule to decode novel call sequences. *Current Biology*, 27(15), 2331–2336. <https://doi.org/10.1016/j.cub.2017.06.031>, e3.
- Suzuki, T. N., & Zuberbühler, K. (2019). Animal syntax. *Current Biology*, 29(14), R669–R671. <https://doi.org/10.1016/j.cub.2019.05.045>
- Townsend, S. W., Engesser, S., Stoll, S., Zuberbühler, K., & Bickel, B. (2018). Compositionality in animals and humans. *PLoS Biology*, 16(8), Article e2006425. <https://doi.org/10.1371/journal.pbio.2006425>
- Tweheyo, M., Lye, K. A., & Weladji, R. B. (2004). Chimpanzee diet and habitat selection in the Budongo forest reserve, Uganda. *Forest Ecology and Management*, 188(1–3), 267–278. <https://doi.org/10.1016/j.foreco.2003.07.028>
- Watson, S. K., Townsend, S. W., & Range, F. (2018). Wolf howls encode both sender and context-specific information. *Animal Behaviour*, 145(10), 59–66. <https://doi.org/10.1016/j.anbehav.2018.09.005>
- White, F. J., & Wrangham, R. W. (1988). Feeding competition and patch size in the chimpanzee species *Pan paniscus* and *Pan troglodytes*. *Beyond Behavior*, 105, 148–164.
- Wrangham, R. W. (1977). *Feeding behaviour of chimpanzees in Gombe National Park, Tanzania*. In T.H. Clutton-Brock, *Primate ecology: Studies of feeding and ranging behaviour in lemurs, monkeys and apes* (pp. 504–538). London, U.K.: Academic Press.

Appendix

Table A1
List of chimpanzees from the Sonso community

Name ID	Sex	Age
Alice	Female	Infant
Anna	Female	Adult
Cyndi	Female	Adult
Deli	Female	Adult
Dembe	Female	Infant
Dora	Female	Adult
Eve	Female	Adult
Faida	Female	Subadult
Faith	Female	Juvenile
Flora	Female	Adult
Frank	Male	Adult
Geoffrey	Male	Juvenile
Gladys	Female	Adult
Goria	Female	Subadult
Hadue	Male	Infant
Harmoni	Female	Juvenile
Harriet	Female	Adult
Hawa	Male	Adult
Heri	Female	Subadult
Irene	Female	Adult
Ishe	Female	Infant
Jacintha	Female	Infant
Jacob	Male	Juvenile
James	Male	Subadult
Janie	Female	Adult
Juliet	Female	Adult
Kaija	Male	Juvenile
Kalema	Female	Adult
Kaqwa	Male	Infant
Kasigwa	Male	Adult
Kaspa	Female	Subadult
Kathy	Female	Subadult
Katia	Female	Adult
Kato	Male	Adult
Kavera	Male	Juvenile
Kefa	Male	Juvenile
Kewayya	Female	Adult
Kigere	Female	Adult
Kirabo	Male	Juvenile
Klauce	Male	Subadult
Kox	Female	Subadult
Kutu	Female	Adult
Kwera	Female	Adult

Table A1 (continued)

Name ID	Sex	Age
Kwezy	Male	Adult
Linda	Female	Adult
Marion	Female	Adult
Mbotella	Male	Juvenile
Melissa	Female	Adult
Muhumuza	Male	Infant
Mukwano	Female	Adult
Musa	Male	Adult
Nambi	Female	Adult
Oakland	Female	Adult
Ozzie	Male	Juvenile
Pascal	Male	Adult
Rafia	Female	Subadult
Ramula	Female	Adult
Roman	Male	Infant
Rose	Female	Adult
Ruhara	Female	Adult
Sharlot	Female	Subadult
Simon	Male	Adult
Squibs	Male	Adult
Tanja	Female	Adult
Twenty	Female	Subadult
Upesi	Female	Adult
Waseme	Female	Adult
Yuliyo	Female	Adult
Zalu	Male	Adult
Zambe	Female	Adult
Zed	Male	Adult
Zefa	Male	Adult

Table A2
Hierarchy calculated based on pant grunt production and Elo-ratings

ID	Elo-rating
Hawa ^a	1776
Musa ^b	1581
Frank	1413
Simon	1093
Squibs	993
Zalu	931
Kato	927
Pascal	888
Zefa	877
Kwezy	560
Zed	494
Kasigwa	411

^a Alpha male.

^b Beta male.

Table A3
Frequencies/h of pant hoots and food calls produced in isolation for adult males

ID	Pant hoot/h	Food call/h
Hawa	0.1	0.1
Musa	0.09	1.48
Frank	0.11	0.33
Simon	0	0
Squibs	0	0
Zalu	0.12	4.24
Kato	0	1
Pascal	0.5	1.17
Zefa	0	5.5
Kwezy	NA	NA
Zed	0.25	2.38
Kasigwa	0	0.8

NA: not available.

Table A4
Multiple distinctive collocation analysis for bigrams

Call type	pbin_b	pbin_c	pbin_fc	pbin_pb	pbin_pg	pbin_ph	pbin_s	pbin_w	SumAbsDev
B	−0.03522	3.819	−0.03522	−0.03522	−0.66555	−0.63691	−0.05319	−0.05319	5.3335
FC	−0.26998	−0.54747	−0.26998	−0.26998	−5.10252	<u>9.49126</u>	−0.40776	−0.40776	16.76671
P	−0.05869	−0.11901	−0.05869	−0.05869	1.9897	−1.06152	−0.08864	−0.08864	3.52358
PB	−0.07043	−0.14282	−0.07043	−0.07043	1.38764	−0.59416	−0.10637	−0.10637	2.54865
PG	0.82476	−0.14282	−0.07043	0.82476	−1.33109	0.79241	−0.10637	−0.10637	4.19901
PH	−0.19955	−0.40465	1.12948	0.43369	3.34544	−3.60916	−0.30139	−0.30139	9.72475
S	0.66569	−0.21423	−0.10565	−0.10565	0.57419	−1.91073	−0.15956	2.34849	6.08419
WB	−0.07043	0.55246	−0.07043	−0.07043	−0.63212	−0.59416	2.93239	−0.10637	5.02879

Columns refer to the first call of bigrams while rows show the second call (ph = pant hoot, fc = food call). Pbins are the log-transformed P values of the contrasted calls (pbin $* > 3 \Rightarrow P < 0.001$, $* > 2 \Rightarrow P < 0.01$, $* > 1.30103 \Rightarrow P < 0.05$). Significant results are shown in bold meaning that there is an attraction between specific calls. The value for the pant–hoot–food call combination is underlined.

Table A5
GLMM indicating the influence of the duration of feeding event (min) on the joining of high-ranking individuals (0/1)

	Estimate	SE	z	P
(Intercept)	−2.317	0.515	−4.496	
Duration	0.006	0.004	1.506	0.132

To rule out the possibility that the reported effect of feeding duration on call combination production could be driven by the fact that high-ranking males are simply more likely to join the longer the feeding event, we ran a GLMM with a binomial error structure. We fitted the joining of a high-ranking individual (0/1) as the binary response variable and the duration of the feeding event (min) as the predictor. The results reported in the table show no statistical relationship between the duration of the feeding event and the probability of high-ranking individuals joining ($P=0.132$).