BRILL

Behaviour 151 (2014) 1091–1120

Behaviour
brill.com/beh

Acoustic structure and variation in mountain and western gorilla close calls: a syntactic approach

Daniela Hedwig [a,*], **Kurt Hammerschmidt** [b], **Roger Mundry** [a], **Martha M. Robbins** [a] and **Christophe Boesch** [a]

[a] Max Planck Institute for Evolutionary Anthropology, Department of Primatology, Deutscher Platz 6, 04103 Leipzig, Germany

[b] German Primate Centre, Cognitive Ethology Lab, Kellerweg 4, 37077 Göttingen, Germany

[*]Corresponding author's e-mail address: hedwig@eva.mpg.de

Accepted 10 January 2014; published online 14 February 2014

Abstract

Our understanding of the functioning of a species' vocal repertoire can be greatly improved by investigating acoustic variation and using objective classification schemes based on acoustic structure. Here we used a syntactic approach to investigate the acoustic structure of the gorilla close distance vocalizations ('close calls'), which remain as yet little understood. We examined 2130 calls of 10 mountain gorillas (*Gorilla beringei beringei*) from Bwindi Impenetrable National Park, Uganda, and 5 western lowland gorillas (*Gorilla gorilla gorilla*) from Bai Hokou, Central African Republic. We segmented calls into units using distinct acoustic features and employed model-based cluster analyses to define the repertoire of unit types. We then examined how unit types were combined into calls. Lastly, we compared unit type use between age–sex classes and the two study groups. We found that the gorilla close calls consist of 5 intergraded acoustic unit types which were flexibly but yet non-randomly concatenated into 159 combinations. Our results are in line with previous quantitative acoustic analyses demonstrating a high degree of acoustic variation in a variety of animal vocal repertoires, particularly close distance vocalizations. Our findings add on to (1) the recent argument that the common practice of describing vocal repertoires as either discrete or graded may be of little value as such distinctions may be driven by human perception and non-quantitative descriptions of vocal repertoires, and (2) recent studies indicating that flexibility in close range social calls can come about through combinatorial systems, which previously have been studied primarily in long distance vocalizations. Furthermore, our study highlights differences in the vocal repertoire of western and mountain gorillas, as expected given differences in environment and social behaviour. Our results offer opportunities for further in-depth studies investigating the function of the gorilla close calls, which will contribute to a more comprehensive understanding of ape vocal communication in general.

Keywords

close calls, vocal repertoire, acoustic analysis, syntax, mountain gorilla, western gorilla.

DOI:10.1163/1568539X-00003175

1. Introduction

Vocal repertoires have commonly been described by classifying vocalizations by ear, visual inspection of spectrograms, or context with subsequent quantitative acoustic descriptions of the identified call types. However, our understanding of the functioning of a species' vocal repertoire can be greatly improved by investigating acoustic variation and using objective classification schemes based on acoustic features. This is because, first, call types identified by human observers may not be meaningful to the animals under investigation since they may recognize an assigned call type as several different ones based on variation within them (May et al., 1989; Stebbins & Sommer, 1992; Fischer, 1998; Slocombe et al., 2009). Moreover, most mammalian repertoires are graded systems exhibiting continuous variation within and between call types (e.g., *Macaca sylvanus*: Hammerschmidt & Fischer, 1998; *Suricata suricatta*: Manser, 2001; *Sus scrofa*: Tallet et al., 2013; *Cercopithecus campbelli*: Keenan et al., 2013), and this acoustic variation can be meaningful as it may correlate with caller identity (e.g., *Pan troglodytes*: Mitani et al., 1996; *Tursiops truncates*: Janik et al., 2006; *Panthera tigris*: Ji et al., 2013), body size (e.g., *Papio hamadryas*: Pfefferle & Fischer, 2003; *Bison bison*: Wyman et al., 2012; *Phascolarctos cinereus*: Charlton et al., 2012), and affect intensity (e.g., *Macaca sylvanus*: Fischer et al., 1995; *Suricata suricatta*: Manser, 2001). Additionally, vocalizations of animal species ranging from birds to primates are composed of acoustic subunits, whereby the meaning of the vocalization is conveyed in the particular arrangement of these units (Marler, 2000; Berwick et al., 2011; ten Cate & Okanoya, 2012). Particularly in primates a vocal signal can act as a modifier of the meaning of another signal when combined (*Saguinus oedipus*: Cleveland & Snowdon, 1982; *Cebus olivaceus*: Robinson, 1984; *Cercopithecus diana*: Zuberbühler, 2002; *Cercopithecus campbelli*: Ouatarra et al., 2009) and signal combinations can impart different information compared to their singularly used components (*Pan troglodytes*: Crockford & Boesch, 2005; *Cercopithecus diana*: Candiotti et al., 2012). Such syntactic variation can encode information about the presence of a predator (e.g., Zuberbühler, 2002; Clarke et al., 2006) and food quality (e.g., *Pan paniscus*: Clay & Zuberbühler, 2012) or on-going behaviour of the caller or external events (e.g., Crockford & Boesch, 2005; Candiotti et al., 2012). Lastly, only a detailed, objective and quantitative assessment of the acoustic structure of a species' vocal repertoire provides the necessary foundation for comparative studies investigating the influence of environmental (e.g., Morton, 1975; Waser & Brown, 1986) and social forces on the evolution of vocal communication systems (e.g., McComb & Semple, 2005; Freeberg et al., 2012) as well as the animals' capacity for vocal learning (e.g., Crockford et al., 2004).

The goal of this study was to provide a quantitative acoustic analysis of the repertoire of close range vocalizations of one group of mountain gorillas and one group of western gorillas, following a syntactic approach. Close range vocalizations or so-called close calls (Harcourt et al., 1986, 1993; Seyfarth et al., 1994; Harcourt & Stewart, 1996, 2001) are characteristic for gorilla vocal behaviour, however, the detailed acoustic structure and function of these calls remain as yet little understood. A detailed investigation of the gorilla close calls will significantly contribute to a more comprehensive understanding of ape vocal communication, which has previously been regarded as affective and restricted to conveying information about emotional states. While an increasing number of studies revealed high degrees of social cognition and flexibility involved in ape vocal behaviour (e.g., *Pan troglodytes*: Crockford & Boesch, 2004; Slocombe & Zuberbühler, 2005; Crockford et al., 2012; Gruber et al., 2013; Schel et al., 2013; *Pan paniscus*: Clay & Zuberbühler, 2011; *Pongo* sp.: Wich et al., 2012; but see Owren et al., 2011), including syntactic vocalizations (e.g., Crockford & Boesch, 2005; Clay & Zuberbühler, 2012), the little we know about gorilla communication stands out as striking. Furthermore, primate vocal communication is dominated by the use of close distance vocalizations, which are a crucial component of primate social behaviour and as such regulate many aspects of group life (e.g., Boinski, 1993; Cheney et al., 1995; Boinski & Campbell, 1996; Silk et al., 2000; Witham et al., 2007). However, research on primate vocalizations, particularly syntactical research, has often focused on long distance vocalizations. Additionally, the genus gorilla is an excellent system for investigating the socioecological driving forces of the evolution of vocal communication systems because gorillas live in a variety of habitats which is expected to lead to variability in their social behaviour (e.g., Robbins, 2010), which in turn should lead to variation in vocal behaviour. However, most of what we know about gorilla vocal behaviour derives from mountain gorillas, which live in a unique high altitude habitat (e.g., Harcourt et al., 1993; but see Salmi et al., 2013 for western gorillas).

The gorilla close calls are a group of intergraded vocalizations of relatively low pitch, ranging from short grunts to longer grumbles and hums (Schaller, 1963; Fossey, 1972; Harcourt et al., 1986, 1993; Seyfarth et al., 1994; Salmi et al., 2013). These calls typically consist of several segments (Harcourt et al., 1993; Salmi et al., 2013) suggesting a syntactic structure. Calls were classified into categories identified by ear in the field with subsequent acoustic analysis of the identified call types (Harcourt et al., 1986, 1993). Two broad classes of vocalizations, subdivided into a total of nine acoustically distinct close call types, were described (Table 1). Syllabled calls were roughly identified as staccato like low pitched grunts of short duration. Non-syllabled calls were described as less sharply articulated low

Table 1.
Overview of close call types and their contexts described for mountain gorillas (MG) and western gorillas (WG).

Call type	Species	Context
Syllabled calls		Grooming, in response to noise, by mothers to infants, towards the end of resting periods (Harcourt & Stewart, 1986; Harcourt et al., 1993; Stewart & Harcourt, 1994)
Single grunt	MG/WG	Feeding, resting, travelling (Salmi et al., 2013)
Double grunt	MG/WG	Individually distinct (Seyfarth et al., 1994); feeding, resting, travelling (Salmi et al., 2013)
2 subtypes	MG	Spontaneous and reply call (Seyfarth et al., 1994)
Triple grunt	MG	Unknown
Inverted grunt	MG	Unknown
Train grunt/whinny*	MG/WG	Mating (Harcourt et al., 1993; Salmi et al., 2013; Watts et al., 1991; Sicotte, 1994)
Non-syllabled calls		As chorus when individuals are feeding and moving close together (Harcourt et al., 1993)
Grumble	MG/WG	More by low ranking as compared to high ranking individuals (Harcourt et al., 1993); feeding, resting, travelling (Salmi et al., 2013)
Hum	MG/WG	Mainly feeding, rarely resting and travelling (Salmi et al., 2013)
High hum/sing	MG/WG	Feeding (Salmi et al., 2013)
Dog whine	MG	Unknown

* Not included in our analysis because acoustically and contextually distinct from other close calls.

pitched growls and high pitched hums of longer duration (Harcourt et al., 1986). However, the authors acknowledged that this classification underestimates acoustic variation and that close calls are intergraded with intermediate forms potentially conveying different information (Fossey, 1972; Harcourt et al., 1986, 1993; Seyfarth et al., 1994; Salmi et al., 2013). Gorilla close calls seem to play an important role in coordinating the interactions between group members as they are usually given when individuals are in close proximity, as part of vocal exchanges, and more while feeding as compared to resting (Harcourt et al., 1986). However, due to their overall seemingly context unspecific nature the detailed function of the gorilla close calls is little understood (Table 1). A recent study of Salmi et al. (2013) identified 6 acoustically distinct close call types in western gorillas (Table 1) and suggests that the use of close calls is more context specific in western gorillas as compared to mountain gorillas as an adaptation to larger group spread in western gorillas.

Here, we conducted an acoustic analysis of close calls of mountain gorillas in Bwindi Impenetrable National Park, Uganda, and western gorillas in Bai Hokou, Central African Republic. Our specific goals were to provide (i) a starting point for future in-depth studies on the informational content and meaning of the gorilla close calls and (ii) the first directly comparable account of the vocal repertoires of western and mountain gorillas. We used a syntactic framework by describing the acoustic structure of close calls as combinations of acoustic units. First, we segmented calls into units and used model-based cluster analyses to classify units into unit types and assess the degree of intergradation between these. Next, we examined how unit types were combined into calls. Lastly, we compared our results between age–sex classes and the two study groups. We expected that units can be classified into potentially intergraded unit types (e.g., Keenan et al., 2013). Furthermore, we expected unit types to be flexibly combined and that their combination follows some regularities which are consistent across individuals (e.g., Cleveland & Snowdon, 1982; Robinson, 1989; Crockford et al., 2004). We also expected some differences between the vocal repertoires of the two groups, reflecting potential adaptations to differences in environment and sociality of the two gorilla species (e.g., Salmi et al., 2013), as well as differences between age–sex classes due to differences in life-history strategies and body size (e.g., Harcourt et al., 1993).

2. Material and methods

2.1. Study sites and study groups

We observed one group of mountain gorillas in Bwindi Impenetrable National Park (0°53′–1°08′N, 29°35′–29°50′E), Uganda. The study site consists of afromontane rainforest (altitude 1160–2600 m) which is characterized by steep hills and a dense understorey of herbaceous vegetation. Data collection focused on the adult individuals (one silverback, five adult females, four blackbacks) of the habituated 'Kyagurilo' group (Robbins, 2008), consisting of 16 animals at the time of the study (two additional juveniles and three infants). One old adult female was excluded from the analysis of vocal behaviour due to her raspy voice and thus barely audible vocalizations. Due to regulations of the Uganda Wildlife Authority observations were restricted to approximately four hours per day. Data collection was conducted usually in the mornings, on 312 days during a total of 12 months from October 2007 through October 2008. A total of 33 h of focal time was collected for each adult individual (range 32.6–34.6). One group of habituated western lowland gorillas was observed at the Bai Hokou study site (2°51′N, 16°28′E), located in the Dzanga-Ndoki National Park, Central African Republic. The habitat is a low altitude mixed-species semi-evergreen rainforest interspersed with areas of *Gilbertiodendron dewevrei* (Caesalpiniaceae) forest. Data were collected in the mornings (7:00–12:00) and/or afternoons (12:00–17:00) on 124 days over a period of eight months from April to November 2009 on all adult individuals (one silverback, three adult females, one blackback) of the 'Makumba' group (Masi et al., 2009), which at the time of the study consisted of 11 group members (three additional juveniles and three infants). A total of 58 h of focal time was collected for each adult individual (range 57–58.5). We considered our study subjects as adults since all the females were parous and since blackbacks (age ca. 11–13 years) are known to sire offspring (Bradley et al., 2005).

2.2. Recordings and acoustic measurements

Audio recordings were conducted during focal animal sampling (Altmann, 1974). Additionally, in order to increase our sample size we opportunistically recorded calls given by individuals in close proximity to the focal animal. Recording distance was on average ca. 5 m. Occasionally we included calls into the analysis that were recorded at a distance of up to 15 m. Vocalizations were recorded at a 48 kHz sampling rate using a Marantz PMD670 digital audio recorder with a Sennheiser ME66 shotgun microphone and a K6 powering module. For each recorded close call we noted the identity of the caller and the call type according to previous classifications of mountain gorilla close calls (Harcourt et al., 1986, 1993). We distinguished between syllabled and non-syllabled calls and differentiated between single, double and triple grunts, depending on the number of grunts, as well as low pitched grumbles and high pitched hums. We used an additional category that was not listed by Harcourt et al. (1993) for non-syllabled calls that contained both low and high pitched components and often comprised several segments, which we labelled as mixed calls (see Figure 1 for examples). Even though Harcourt et al. (1993) distinguished between more call types, we used these broader categories as they allowed the most unambiguous classification of calls by ear in the field. We also identified train grunts but we excluded these from further analysis as they are acoustically as well as contextually highly distinct (e.g., Watts, 1991; Harcourt et al., 1993; Sicotte, 1994).

We recorded a total of 9476 close calls. In order to select high quality vocalizations, recordings were downsampled to 2000 Hz, and spectrograms were produced using Avisoft SASlab Pro Version 5.1.23 (R. Specht, Berlin, Germany). Based on call/background noise ratio we selected a total of 2130 calls from 15 individuals for acoustic analysis (Figure 2). Our analysis included 943 double grunts, 279 single grunts, 4 triple grunts, 437 grumbles, 183 hums and 284 mixed close calls (Table 2).

Based on visual inspection of spectrograms, aiming to fully capture the combinatorial properties while obtaining a similar representation of the gorilla close calls compared to previous studies (Harcourt et al., 1993; Salmi et al., 2013), we defined calls as bouts of vocal utterances separated by periods of silence of more than 2 s and divided calls into units based on two criteria: (1) the occurrence of periods of silence of a maximum of 2 s (min $= 0.01$ s) between them (Figure 1a–e, h, i), (2) sudden spectral shifts, indicating abrupt changes between high pitched tonal hums and low pitched grunts, which were characteristic for many mixed calls (Figure 1j, k). Subsequently, units were categorized into atonal and tonal due to the presence/absence of harmonic frequency bands. This was necessary since the presence or absence of harmonics determined the sets of acoustic parameters used to quantify these units. Pauses and spectral shifts have been shown to be distinct and functionally relevant for classifying acoustic units in a number of similar studies

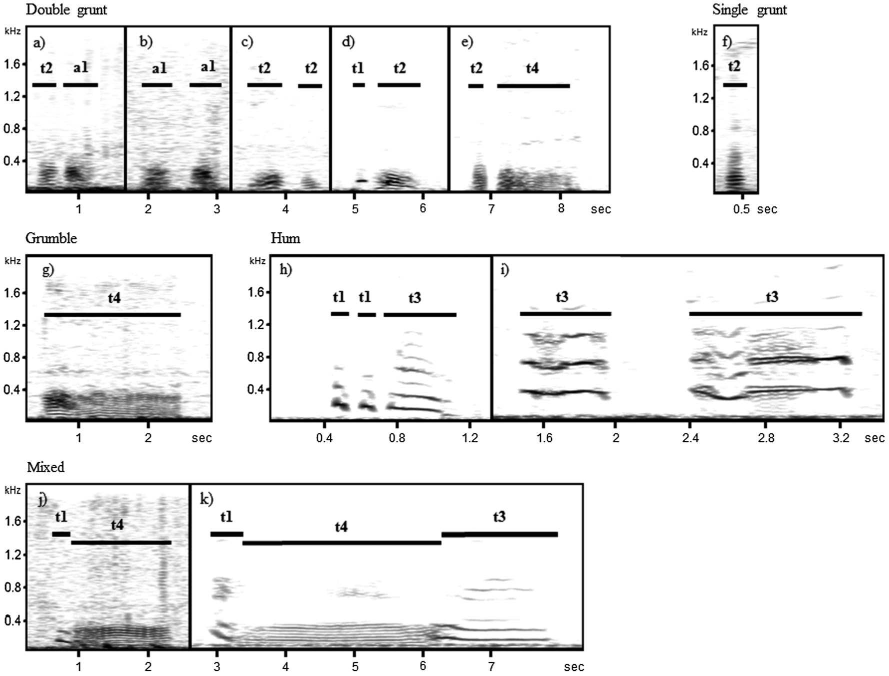


Figure 1. Spectrograms of representative gorilla close calls recorded in this study, their subdivision into units and categorization into unit types. Spectrograms a–f illustrate typical examples of syllabled calls: (a–e) double grunts, (f) single grunt. Spectrograms g–k illustrate non-syllabled calls: (g) grumble; (h, i) hums; (j, k) mixed calls. Calls were subdivided into units (indicated by black lines) based on the occurrence of periods of silence of less than 2 s duration (a–e, h, i) or abrupt changes in the distribution of energy (j, k). The units were categorized as atonal or tonal according to the presence/absence of harmonic frequency bands. Indicated above the lines are the unit type each unit was assigned to via cluster analysis: a1, atonal grunts; t1, short hums; t2, short tonal grunts; t3, long hums; t4, grumbles.

(*Homo sapiens*: Furui, 1986; *Sturnus vulgaris*: Eens et al., 1989; Franz & Goller, 2002; *Orcinus orca*: Shapiro et al., 2011; *Mungos mungo*: Jansen et al., 2012).

For acoustic analysis we chose a set of 4 parameters for the tonal and 6 parameters for atonal units that broadly describe frequency-related and temporal characteristics of the units (Table 3). It is unlikely that variation in recording distance strongly influenced the acoustic measurements as it was on average 5 m. Furthermore, the acoustic parameters chosen revealed a high accuracy under our recording distances (Maciej et al., 2011). Additionally, we were in visual contact with the caller when audio recording, ruling out

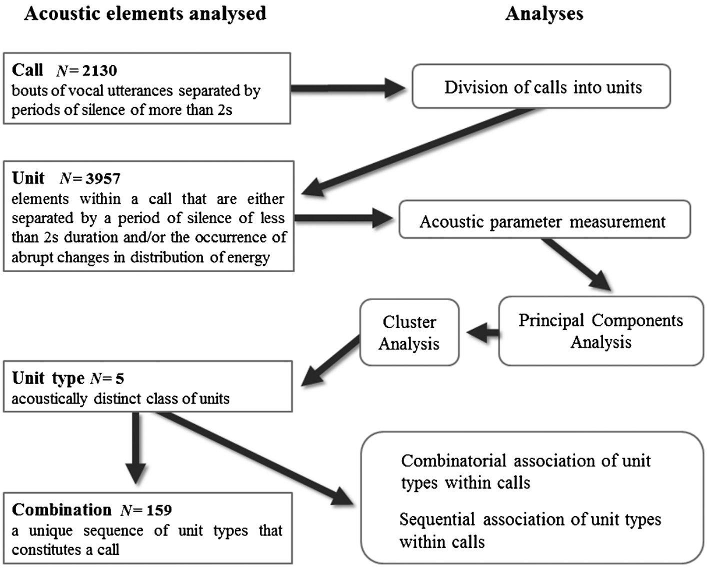


Figure 2. Operational definitions and schematic overview of analyses used in this study.

severe acoustic distortion of calls propagating through dense vegetation (e.g., Bradbury & Vehrencamp, 2011).

Durations of units were measured manually from spectrograms with a 20 Hz frequency and 1 ms temporal resolution using Avisoft SASlab Pro. Frequency-related parameters were measured semi-automatically on spectrograms of 4 Hz frequency and 4 ms temporal resolution using the software LMA 2012 developed by K. Hammerschmidt (Hammerschmidt, 1990; Fischer et al., 2013). For atonal units we measured acoustic parameters describing the peak frequency, which is the frequency of the highest amplitude in a given unit. We limited these measurements to a range between 70 and 400 Hz, roughly focusing on the first dominant energy band while avoiding background noise that could confound our measurements. For tonal units we extracted parameters describing the fundamental frequency. Due to the noisy character of gorilla grunts (Harcourt et al., 1993; Salmi et al., 2013) or sudden frequency modulations LMA was sometimes not able to perform these measurements. In these cases, parameters describing the fundamental frequency were manually measured in Avisoft SASlab Pro. Reliability of these manual measurements was ascertained through comparing the results of manual with automatic measurements on a subset of tonal units (Spearman correlation; $\rho > 0.92$, $N = 55$, $p < 0.001$). For all further analyses,

Table 2.
Overview of sample sizes for close call types, unit types, and combinations found in close calls of the individuals of the mountain gorilla and western gorilla group.

	Mountain gorilla													Western gorilla							Total
	AF					Total AF	BB				Total BB	SB	Total mountain gorilla	AF			Total AF	BB	SB	Total western gorilla	
	BY	KK	MG	ST	TN		BZ	FZ	MR	SK		RC		BE	MAL	MPB		KU	MK		
Close call types																					
Double grunts	55	70	73	29	14		94	92	54	40		113	634	45	138	24		45	57	309	943
Single grunts	5	30	7	18	0		25	12	12	6		6	121	26	25	56		24	27	158	279
Triple grunts	0	3	0	0	0		0	0	1	0		0	4	0	0	0		0	0	0	4
Grumbles	4	37	27	17	0		1	19	1	1		5	112	53	102	39		20	111	325	437
Hums	8	2	13	3	9		35	8	10	19		0	107	5	0	1		68	2	76	183
Mixed	18	14	34	23	4		57	40	21	11		0	222	4	1	1		31	25	62	284
No. calls total	90	156	154	90	27	517	212	171	99	77	559	124	1200	133	266	121	520	188	222	930	2130
No. units total	178	269	315	160	53	975	438	368	214	148	1168	238	2381	186	418	184	788	454	334	1576	3957
Unit types																					
Atonal grunt a1	74	14	18	21	2		126	59	39	18		155	526	57	53	40		43	52	245	771
Short hum t1	12	17	31	21	23		86	51	47	25		5	318	5	2	4		194	18	223	541
Tonal grunt t2	37	74	108	23	7		48	65	40	42		34	478	36	204	80		89	44	453	931
Long hum t3	32	19	54	21	11		89	51	25	32		6	340	5	0	1		124	23	153	493
Grumble t4	23	145	104	74	10		89	142	63	31		38	719	83	159	59		4	197	502	1221
Unit types total	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5
Sampling coverage	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.995	1.00	1	1	0.999	0.998
No. of unique combinations	23	25	32	15	11	53	41	39	31	22	80	12	100	17	17	21	31	63	27	97	159
Sampling coverage	0.91	0.92	0.90	0.92	0.78	0.89	0.89	0.85	0.82	0.83	0.85	0.98	0.88	0.95	0.97	0.91	0.94	0.77	0.94	0.91	0.89

AF, adult female; BB, blackback; SB, silverback; BY, KK, MG, ST, TN, BZ, FZ, MR, SK, RC, BE, MAL, MPB, KU, MK denote individual gorillas.

Table 3.
Acoustic parameters measured for tonal and atonal vocal units and their transformations.

Parameter	Description	Measured for	Transformation
F0max	Maximum fundamental frequency (Hz)	tonal units	Log
F0min	Minimum fundamental frequency (Hz)	tonal units	Log
F0mean	Mean fundamental frequency (Hz)	tonal units	Log
Duration	Duration (s)	tonal/atonal units	Log
Pfmax	Maximum peak frequency across time segments (Hz)	atonal units	None
Pfmaxloc	Location of the maximum peak frequency (location/duration)	atonal units	None
Pfmaxamp	Peak frequency at the maximum amplitude across time segments (Hz)	atonal units	Log
Pfminamp	Peak frequency at the minimum amplitude across time segments (Hz)	atonal units	max(Pfminamp) − (1/Pfminamp)
Pfmaxdiff	Maximum absolute difference between peak frequency values of successive time segments (Hz)	atonal units	square root

Peak frequency is the frequency of the highest amplitude.

acoustic parameters were transformed to achieve more symmetrical distributions and then z-transformed to a mean of zero and a standard deviation of one (Table 3).

2.3. *Statistical analysis*

2.3.1. *Unit type repertoire*

To classify units into unit types and describe the acoustic structure of the unit type repertoire we combined Cluster Analysis with a subsequent assessment of the discreteness of cluster solutions, aiming to choose a classification of unit types that represents the most discrete cluster solution. Because many of the acoustic parameters were clearly correlated to each other, we first performed Principal Components Analyses (PCA) using the R package Psych (Revelle, 2007) (for details, see section A1 of the Appendix, which is part of the online version of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1568539x). We performed Cluster Analyses on the resulting principal components using the R package MClust which provides functions for model-based clustering (Fraley et al., 2009). Cluster solutions were calculated through employing a maximum likelihood method and using the Bayesian Information criterion (BIC: Schwarz, 1978). We measured the discreteness of cluster solutions by calculating silhouette coefficients ($S(c_n)$), which quantify the cohesion within and separation between clusters. A $S(c_n) < 0.25$ indicates strong intergradation, $0.25 \leqslant S(c_n) < 0.5$ weak clustering, $0.5 \leqslant S(c_n) < 0.7$ medium and $S(c_n) \geqslant 0.7$ strong clustering (Rousseeuw, 1987) (for further details on the cluster analyses, see section A2 of the Appendix, which is part of the online version of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1568539x). We further validated the results of the cluster analyses by testing whether clusters emerged through random regrouping of units or instead constituted subclusters of larger clusters (see section A2). We also evaluated how consistent results of all further analyses were across different cluster solutions. Therefore, while the focus was on the classification based on the cluster solution that revealed the most discrete clusters (4 tonal and 1 atonal unit type), we repeated all analyses using the classification of 8 tonal and 7 atonal unit types (detailed results for this can be found in section A3 of the Appendix, which is part of the online version of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1568539x).

2.3.2. Repertoire of combinations

We described each call as a sequence of unit types and defined unique sequences of unit types as combinations (Figures 1, 2). We investigated the similarity between individual combination repertoires by calculating the Dice index of repertoire overlap as $D = (2R_{ij}/(R_i + R_j))$ for pairs of individuals (Cairns & Schwager, 1987), where R_{ij} is the number of combinations that are part of both individuals' repertoires, R_i and R_j represent the number of combinations in the repertoires of the two individuals i and j. A $D = 0$ indicates no overlap between the two individuals' repertoires, whereas $D = 1$ indicates identical repertoires. Based on this, we assessed the similarity in combination repertoires between groups and age–sex classes. Therefore, we ran Mantel-like permutation tests (Sokal & Rohlf, 1995) using one matrix depicting the Dice indices of repertoire overlap for each pair of individuals and another matrix depicting whether individuals belonged to the same group or age–sex class, respectively. Due to small numbers of individuals in each age–sex class, all group or age–sex class comparisons focused on adult females of the mountain and western gorilla group or on the blackbacks and adult females of the mountain gorilla group.

2.3.3. Combinatorial association of unit types

We investigated how frequently two unit types occurred together in the same call. For this we calculated the Dice index of combinatorial association for each pair of unit types (per individual) as above. Here R_{ij} is the number of calls including both unit types and R_i and R_j represent the number of calls including the unit types i and j, respectively. To compare the combinatorial associations between the two study groups and age–sex classes we ran Mann–Whitney U-tests on average Dice indices per individual.

2.3.4. Sequential association of unit types

We investigated (1) how frequently any two unit types followed each other successively within calls by calculating the Dice index of sequential association (as above) for each pair of unit types (per individual). Here R_{ij} is the number of times in which unit type j followed in immediate succession of unit type i, and R_i and R_j are the numbers of times unit type i preceded any unit type and j followed any unit type, respectively. (2) We measured how random the sequential order of unit types within each individual's calls was. We compiled individual transition matrices depicting the frequency of each unit type to follow each other unit type (or being repeated) and calculated individual contingency coefficients as $C = \sqrt{\chi^2 \times m/((\chi^2 + N) \times (m - 1))}$. Here, χ^2 is the chi-square statistic of the transition matrix, N is the total number of transitions in the matrix and m is the minimum of the number of its rows and columns. A contingency coefficient close to 1 indicates that the sequential order of unit types was highly predictable. (3) We investigated how consistent the sequential ordering of unit types was across individuals, groups and age–sex classes. We assessed the similarity in sequential order of unit types for each pair of individuals with Spearman correlation coefficients calculated with Mantel-like permutation tests on pairs of individual transition matrices, whereby we considered only those unit types that were present in both individuals to be compared. Based on this, we compared the similarity in the sequential order of unit types between groups and age–sex classes using Mantel-like permutation tests as described above. If not specified otherwise, results are indicated as average values across individuals and ranges are given in brackets. All analyses were done in R, version 2.15.0.

2.3.5. Sampling effort

To investigate if our sampling effort sufficed to assess the full repertoire of unit types and combinations we calculated the sample coverage Q as the probability that an additionally recorded unit or call would be of a unit type or combination already documented in our sample (Fagen & Goldman, 1977). The sample coverage is measured as $Q = 1 - (N_1/I)$, where N_1 is the number of unit types or combinations recorded only once, and I is the total number of units or calls recorded. If Q is very close to 1, it means that the probability of observing a novel unit type or combination in an additional recording is low and that we largely captured the repertoire of an individual. Our sampling was sufficient to compile a complete repertoire of unit types, indicated by a sampling coverage of 1 for all but one individual (MPB). We were close to having sampled the complete repertoire of combinations for both the mountain and western gorilla individuals (mountain gorilla: average sample coverage $Q = 0.88$ (0.78–0.98, $N = 10$ individuals), western gorilla: average sample coverage $Q = 0.91$ (0.78–0.97, $N = 5$ individuals; Table 2)).

3. Results

3.1. Unit type repertoire

We found that the unit type repertoire was highly intergraded but could be divided into 5 unit types. We categorized all atonal units into one unit type (a1) because silhouette coefficients consistently <0.2 in all cluster solutions indicated that the identified clusters were highly intergraded. A silhouette coefficient >0.3 indicated a weak but consistent clustering across cluster solutions for tonal units. We classified the tonal units into four unit types representing the most discrete cluster solutions (silhouette coefficient $= 0.48$) (see section A2). We assigned short labels to these 5 unit types (a1, t1, t2, t3, t4) convenient for representing calls as combinations (Figure 1). Additionally, we labelled each unit type according to previous studies on gorilla communication (Harcourt et al., 1993; Salmi et al., 2013). Atonal grunts (a1) were on average short, noisy grunts that however could be up to 2 s long. Short hums (t1) and long hums (t3) are clearly tonal unit types with a similarly high frequency range, which were mainly distinguishable by their duration. Additionally, tonal grunts (t2) were usually under a second short grunts, while grumbles (t4) were up to 5 s long (Table 4, Figure 3). The 5 unit types were found in all individuals but one western gorilla female (Table 2).

Table 4.
Overview of values of acoustic parameters characterising the atonal and tonal unit types.

	Duration	Pfmax	Pfmaxloc	Pfmaxdiff	F0mean	F0max
Atonal unit type						
Atonal grunts (a1)	0.32 (0.03–2.17)	184 (74–384)	0.52 (0–1)	63 (0–290)		
Tonal unit types						
Short hums (t1)	0.16 (0.04–0.35)				160 (70–542)	188 (70–570)
Tonal grunts (t2)	0.22 (0.03–0.63)				40 (13–82)	42 (15–109)
Long hums (t3)	1.08 (0.34–4.03)				161 (60–450)	208 (31–464)
Grumbles (t4)	1.34 (0.34–5.01)				31 (11–65)	35 (11–82)

Values are averages and ranges are indicated in parentheses.

3.2. Combinatorial and sequential association between unit types

Overall, the 5 unit types were arranged into 159 combinations (Table 2). We recorded the different combinations on average 13 times, however, the majority was only registered once (number of observations per combination: median $= 1$, range 1–479, 1st quantile $= 1$, 3rd quantile $= 3$). All unit types were used both singularly and combined with other unit types within calls. Also, all unit types were observed together and following successively within the same call. Accordingly, a low average Dice index of combinatorial association ($D = 0.18$ across individuals, 0.08–0.26, $N = 15$ individuals) and sequential association ($D = 0.16$, 0.13–0.19) indicated that the unit types were flexibly arranged into calls. Nevertheless, the sequential arrangement of unit types was non-random and followed some regularities as suggested by an average contingency coefficient of 0.65 across the individual transition matrices (0.21–0.85, $N = 15$ individuals). Furthermore, individuals arranged unit types into similar sequences since we found a significant ($p \leqslant 0.05$) positive correlation for 39% of the 105 comparisons between transition matrices of pairs of individuals (average correlation coefficient 0.38, 0.40–0.91; Mantel-like permutation test). Some unit types occurred more frequently within the same call and in immediate succession than others as indicated by varying Dice indices of association. While the combinatorial association strength between unit types ranged from average minimum Dice

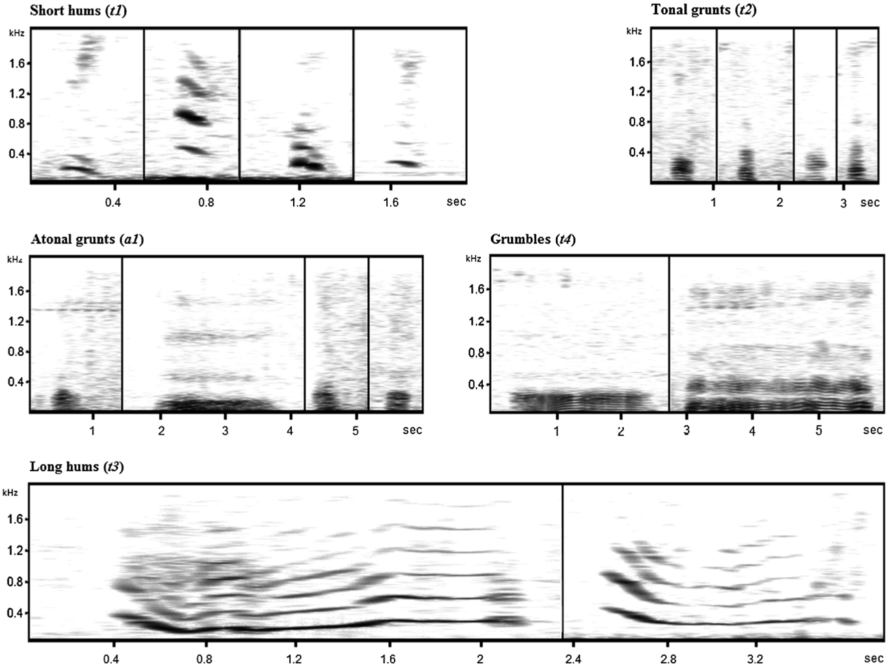


Figure 3. Spectrograms of the 5 unit types. Short hums (t1) were usually under 1-s short, tonal, high pitched units. Tonal grunts (t2) were short, low pitched grunts. Atonal grunts (a1) were on average short, noisy grunts that, however, could be up to 2 s long. Grumbles (t4) were up to 5 s long and low pitched. Long hums (t3) were long tonal, high pitched units.

index of 0.05 (0.01–0.15, $N = 15$ individuals) to an average maximum of 0.47 (0.26–0.58), the sequential association strength ranged from an average minimum Dice index of 0.04 (0.02–0.1) to an average maximum of 0.63 (0.49–1). Notably, short grunts (t2) and grumbles (t4), as well as short hums (t1) and grumbles (t4) were particularly often combined, whereby t2 and t1 usually occurred before t4 within calls (Figures 1, 4). Also long hums (t3) and short hums (t1) were often combined, however, their sequential order was not strongly determined. However, if t3 and a grumble (t4) occurred together in the same call t4 usually preceded t3 (Figures 1, 5).

3.3. *Age–sex class comparisons in mountain gorillas*

The combination repertoires of the adult females and blackbacks were considerably larger than the silverback's repertoire. Congruently, while adult females and blackbacks also used short hums (t1) and long hums (t3) frequently, we mainly recorded the low pitched unit types atonal grunts (a1),

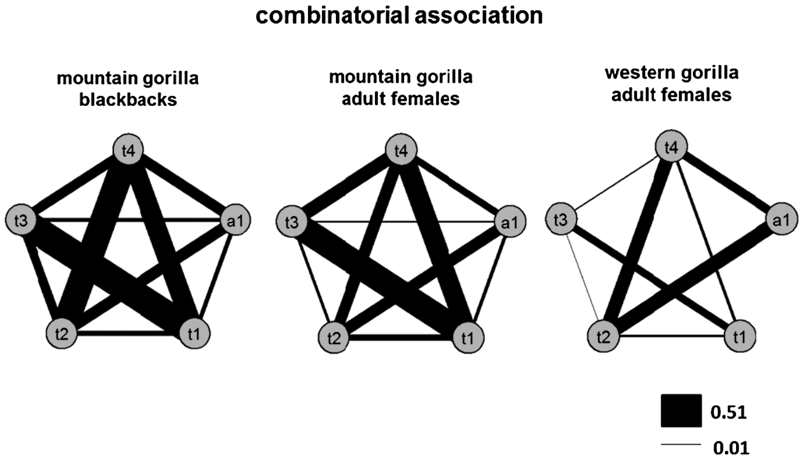


Figure 4. Combinatorial associations between unit types in both gorilla groups. Lines indicate the average Dice index of combinatorial association as a measure of how frequently any two unit types were observed together in the same call. Average values were calculated from the individuals' Dice indices for a given pair of unit types. The legend shows the thickness of the lines corresponding to the minimum and maximum average Dice index.

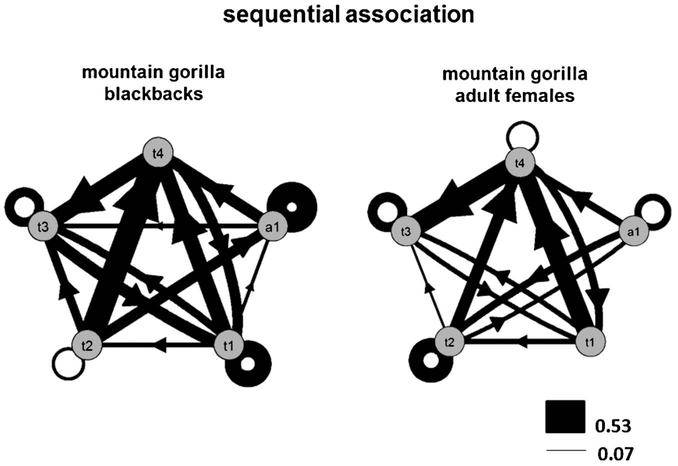


Figure 5. Sequential associations between unit types in the blackbacks and adult females of the mountain gorilla group. Lines indicate the average Dice index for sequential association as a measure of how frequently any two unit types followed one another in immediate succession. Average values were calculated from the individuals' Dice indices for a given pair of unit types. The legend depicts the thickness of the lines corresponding to the minimum and maximum average Dice index. For reasons of clarity illustrations include only transitions with a Dice index > 0.05.

tonal grunts (t2) and grumbles (t4) for the silverback (Table 2). The combination repertoires of females and blackbacks largely overlapped, as the Dice coefficients of repertoire overlap within age–sex class dyads were not significantly larger than those of between age–sex class dyads (permutation test, $N = 4$ blackbacks, 5 adult females, $p = 0.25$). However, we detected differences in how frequently certain unit types were combined and sequentially arranged because the blackbacks tended to show stronger combinatorial associations compared to adult females (exact Mann–Whitney U-test, $U = 3$, $N = 4$ blackbacks, 5 adult females, $p = 0.06$) and average correlation coefficients of within age–sex class dyads (blackbacks: 0.64 (0.55–0.79), $N = 4$; adult females: 0.38 (0.11–0.78), $N = 5$) were significantly larger than those of between age–sex class dyads (0.38 (0.06–0.67); permutation test, $p = 0.014$). In particular, tonal grunts (t2) and the grumble (t4) were more frequently combined and sequentially arranged by the blackbacks compared to the females. Furthermore, the blackbacks used atonal grunts (a1) and short hums (t1) more frequently in immediate succession compared to the females, whereas the females immediately repeated atonal grunts (t2) more frequently as compared to the blackbacks (Figures 1, 4, 5).

3.4. Group comparison

Overall, the mountain and western gorilla groups showed similar combination repertoire sizes, and individuals of both groups shared a similar proportion of combinations (mountain gorilla: average Dice index of repertoire overlap = 0.45 (0.20–0.63), three combinations shared by all individuals; western gorillas: 0.42 (0.19–0.65), seven combinations shared by all individuals). Overall 38 combinations (24%) were observed in both groups, indicating some differences between their repertoires.

The blackbacks of both groups had larger combination repertoires compared to other individuals. The combination repertoire of the western gorilla females was considerably smaller as compared to the silverback and they rarely used short hums (t1) and long hums (t3). In contrast, in the mountain gorilla group the females' repertoires were larger than the silverback's and they frequently used short hums (t1) and long hums (t3) (Table 2). Congruently, the similarity between combination repertoires of two given females tended to be larger when they were of the same group (adult females mountain gorilla: 0.46 (0.36–0.56), $N = 5$; adult females western gorilla group: 0.62 (0.58–0.65), $N = 3$) as compared to when they belonged to different groups (0.43 (0.19–0.57); permutation test, $p = 0.06$). Furthermore, mountain gorilla females showed stronger combinatorial associations between unit types than the western gorilla females ($U = 15$, $N = 5$ adult females mountain gorilla, 3 adult females western gorilla, $p = 0.04$). In particular, the mountain gorilla females combined t3 and t1, t3 and t4, as well as t1 and t4 more frequently as compared to the western gorilla females (Figures 1, 4). However, they arranged unit types into similar sequences since correlation coefficients of within-group female dyads did not significantly differ from those of between-group dyads (permutation test, $N = 5$ adult mountain gorilla females, 3 adult western gorilla females, $p = 0.125$).

3.5. *Comparison with previous classification of close calls*

The non-syllabled calls were more diverse compared to the syllabled calls as they comprised a larger number of combinations, a larger percentage of combinations that we recorded only once, and a more variable number of units within calls. While the combinations recorded among the non-syllabled calls consisted to a much larger proportion of long hums (t3) and short hums (t1), the syllabled calls consisted mainly of atonal grunts (a1) and tonal grunts (t2) (Table 5). The five most frequently observed combinations were the grumble (t4) used singularly ($N = 479$ observations) and different double grunts (t2_t4 ($N = 271$), a1_a1 ($N = 166$), a1_t4 ($N = 144$), t2_t2 ($N = 129$)) (see Figure 1 for examples). Hums and mixed non-syllabled calls occurred considerably less frequently in western gorilla females compared to mountain gorilla females. We frequently observed mixed non-syllabled calls in the western gorilla silverback, while the mountain gorilla silverback's repertoire consisted largely of double grunts (Table 2).

4. Discussion

In this study we used a syntactic approach to quantitatively analyse the acoustic structure of the gorilla close call repertoire. We found that the close calls contain 5 intergraded types of vocal units, which were high pitched short hums (t1) and long hums (t3), low pitched short grunts (t2) and longer grumbles (t4) as well as low pitched atonal grunts (a1). All 5 unit types were found in 14 of 15 individuals sampled and arranged into 159 combinations. This large acoustic variability among the gorilla close calls seems to be based on the highly flexible arrangement of unit types, as indicated by the relatively

Table 5.
Overview of number of units, unit types and combinations found among the close call types identified in the field.

	Syllabled			Non-syllabled		
	Double grunt	Single grunt	Triple grunt	Grumble	Hum	Mixed
No. of calls	943	279	4	437	183	284
No. of combinations	18	5	3	27	42	90
% combinations observed once	22	0	67	70	62	70
No. of units (mean)	2	1	3	1.2	2.2	3.1
Min	2	1	3	1	1	2
Max	2	1	3	8	8	12
% unit types						
a1	34	21	17	6	0	4
t1	3	4	0	0	44	34
t2	39	34	8	6	4	8
t3	1	2	0	4	52	27
t4	23	39	75	84	0	27

low Dice indices of combinatorial and sequential association between unit types. However, the high contingency coefficients that characterized individual transition matrices show that individuals sequentially arranged unit types in a non-random fashion.

While the large acoustic variation found may partly be inherently due to our subdivision of calls, our results are unlikely to merely represent artefacts of our methodology for several reasons. (i) Despite the graded nature of the unit types our classification was not random, but rather reflects an inherent acoustic structure in which larger clusters split up into smaller sub-clusters (see section A2 of the Appendix, which is part of the online version of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1568539x). (ii) Analyses based on different classifications produced similar patterns with regard to combinatorial and sequential associations between unit types (i.e., flexible combination but non-random sequential order) (see section A3 of the Appendix, which is part of the online version of this journal, which can be accessed via http://booksandjournals.brillonline.com/content/journals/1568539x). (iii) Individuals showed a considerable similarity between their unit type repertoires and how unit types were combined and arranged.

Our finding of strong intergradation between unit types adds to the increasing number of quantitative acoustic investigations characterizing mammalian vocal repertoires as graded systems (e.g., Hammerschmidt & Fischer, 1998; Manser, 2001; Tallet et al., 2013; Keenan et al., 2013). In particular, our results expand on a study demonstrating acoustic intergradation between concatenated call types of the Diana monkey alarm call system, a previous text book example of a discrete communication system (Keenan et al., 2013). As such, our result substantiates the recent argument that the common practice of describing vocal repertoires as either discrete or graded may be of little value as such distinctions may largely be driven by human perception and non-quantitative descriptions of vocal repertoires (e.g., Keenan et al., 2013; Suzuki, 2014). It has been put forth that the variation within and between call types may provide additional subtle information for recipients about the internal state or external events (Manser, 2001; Keenan et al., 2013). However, on the other hand it is possible that individuals may perceive continuous acoustic variation categorically, as shown in a variety of animal species (e.g., *Homo sapiens*: Abramson & Lisker, 1970; *Chinchilla* sp.: Kuhl & Miller, 1976; *Melospiza georgiana*: Nelson & Marler, 1989; *Teleogryllus oceanicus*: Wyttenbach et al., 1996; *Macaca sylvanus*: Fischer, 1998; *Physalaemus pustulosus*: Baugh et al., 2008). Ultimately, playback experiments using the habituation–dishabituation paradigm (e.g., Snowdon & Pola, 1978; Fischer, 1998) can be a fruitful approach to test whether the unit types established through our acoustic analysis correspond to units with distinct meanings in the perception of the gorillas.

Overall, our results are in line with other studies demonstrating a high degree of acoustic variation generated through a flexible, yet non-random combination of acoustic units in primates (e.g., Cleveland & Snowdon, 1982; Robinson, 1984; Mitani & Marler, 1989). The large acoustic variation within the gorilla close calls is consistent with other studies showing particularly high acoustic variability in calls that are given as part of vocal exchanges or that are associated with affiliative behaviours (*Cebuella pygmaea*: Elowson & Snowdon, 1994; *Macaca fuscata*: Sugiura, 2007; *Cercopithecus campbelli*: Lemasson & Hausberger, 2011; *Cercocebus torquatus*: Bouchet et al., 2012). Such flexibility in social calls can be beneficial in the face of variable influences of numerous social factors as compared to context specific calls that target the whole group (e.g., alarm calls; Snowdon, 1997; Bouchet et al., 2012). Therefore, our results support recent studies indicating that flexibility in close range social calls can come about through combinatorial systems (e.g., Candiotti et al., 2012).

The essential question our results provoke is how and what information can be encoded in the syntactic system presenting here. We here suggest two modes of how syntactic variation may be associated to informational variation in gorilla close calls. (1) We observed that all of the 5 unit types could be combined within close calls. Gorillas may retrieve different information depending on the co-occurrence of specific unit types, as found in other primates (e.g., Cleveland & Snowdon, 1982; Robinson, 1984; Zuberbühler, 2002; Ouatarra et al., 2009). More specifically, all of the unit types identified here were used in combination with other unit types but also singularly, and signal combinations may convey different information compared to their singularly used components (e.g., Cleveland & Snowdon, 1982; Robinson, 1984; Crockford & Boesch, 2005; Candiotti et al., 2012). (2) We found that the ordering of unit types was flexible, yet some unit types were more likely to occur in a specific order than in the reversed order. The gorillas may decode specific information from the ordering of unit types, as suggested for other species (Cleveland & Snowdon, 1982; Robinson, 1984; Mitani & Marler, 1989; Crockford et al., 2005; Arnold & Zuberbühler, 2008). However, the strikingly large number of different combinations and particularly the fact that many combinations were observed very infrequently suggests that combinations might not fundamentally differ in their informational content and that not all combinatorial variation may be meaningful for recipients, as shown in putty nosed monkey alarm calls (Arnold & Zuberbühler, 2012). Similarly large combination repertoires have been found in chimpanzees and Japanese great tits (Crockford et al., 2005; Suzuki, 2013). In the highly variable great tit mobbing calls different sets of combinations or specific note types are specific for certain predators while the variation within these sets of combinations may reflect the threat level (Suzuki, 2013). Congruently, gorillas may decode contextual information from the occurrence of certain unit types or their position within a call, while repetitions of unit types within a call might be related to the arousal level of the caller and reflect the urgency of the information transfer. We suggest that future studies explore the communicative significance of the combinations identified here on different structural levels. For example, contextual correlates of variation in the syntactic structure of the gorilla close calls can be analysed with regard to the mere presence or absence of specific unit types, the number of repetitions of a specific unit type, as well as variation in the sequential arrangements of unit types.

Our results match previous descriptions of the gorilla close calls, yet our more detailed quantification of the acoustic variation can serve as a starting point for future studies investigating the functioning of the gorilla close call system. The syllabled calls described as ‘staccato like grunts of short duration’ congruently consisted of short and low pitched unit types a1, t2 and t4, whereas we recorded high pitched and longer unit types t1, t3 and t4 predominantly for the non-syllabled calls defined as ‘less sharply articulated growls and hums of longer duration’ (Harcourt et al., 1986) (Figure 1). Previously, the syllabled calls had been suggested to be of most interest for further studies as these calls were suspected to show most acoustic variation (Harcourt et al., 1986). In line with this, we identified 18 different combinations in what has previously been termed a double grunt comprising two different subtypes (Seyfarth et al., 1994). However, our results suggest the non-syllabled calls and particularly the mixed calls and hums (Figure 1) to be especially intriguing since these were the most variable calls in terms of the number of combinations they comprised (Table 5). Particularly, since these calls seem to mainly occur in a feeding context (D.H., pers. obs.), their acoustic variation could be related to the quality of food discovered by the caller as found in bonobos, *Pan paniscus* (Clay & Zuberbühler, 2011).

Our results, combined with the findings of previous studies, indicate that variation in the environment and sociality may have led to differences in the vocal repertoires of the two gorilla species, as found in other closely related species (*Macaca* sp.: Hohmann, 1989; *Marmotta* sp.: Blumstein & Armitage, 1997; *Papio* sp.: Gustison et al., 2012; *Pongo pygmaeus*: Wich et al., 2012). In accordance with previous studies, we found that in both the mountain and western gorilla group, double grunts and grumbles were the most common vocalizations (Harcourt et al., 1993; Salmi et al., 2013). However, we infrequently sampled calls comprising high pitched unit types t1 and t3, such as hums and mixed non-syllabled calls, from the western gorilla females. Congruently, the hums identified by Salmi et al. (2013) in western gorillas seem to be of lower pitch compared to the mountain gorilla hums (Harcourt et al., 1993). Furthermore, as observed in previous studies, the repertoire of the mountain gorilla silverback was considerably smaller and less diverse than that of other age–sex classes, with a predominant use of double grunts (Harcourt et al., 1993), which, however, was in contrast to the silverback western gorilla whose repertoire was considerably more diverse than that of his females. However, caution should be taken when inferring any species differences from our comparison based on two groups and a small number of individuals. Nevertheless, our findings provide a basis for future studies of causal relationships between socioecological and vocal variation in the two species of gorillas. We specifically suggest comparative investigations of contextual correlates of the syntactic variation in both species aiming to identify functionally equivalent calls or potential functional shifts in acoustically similar calls as adaptations to differences related to the two species' habitats (e.g., Nkurunungi et al., 2005) or social behaviour (e.g., Doran & McNeilage, 2001).

In sum, we here provided a detailed quantitative acoustic analysis of the close calls of mountain and western gorilla individuals. We demonstrated that the gorilla close calls consist of 5 intergraded acoustic unit types which were flexibly but yet non-randomly concatenated into 159 combinations. Such a high degree of potentially meaningful variation within the gorilla close call repertoire is in line with the results of previous quantitative analyses of acoustic variation within mammalian vocal repertoires, particularly close distance vocalizations. Further investigations on the communicative significance of the unit types and combinations identified here offer opportunities for further in-depth studies on the function of the as yet little understood gorilla close calls, which will contribute to a more comprehensive understanding of ape vocal communication in general.

Acknowledgements

We thank the Uganda Wildlife Authority and the Uganda National Council for Science and Technology for permission to conduct our research in Bwindi Impenetrable National Park and the Ministry of Education and Water and Forests of the government of CAR for permission to work in Dzanga-Ndoki National Park. We also thank the Institute of Tropical Forest Conservation in Uganda, as well as the Dzanga Sangha Project in CAR, for logistical support. We are grateful for useful comments on the manuscript by Catherine Crockford and two anonymous reviewers. Special thanks go to the local

Ba'Aka trackers of the Primate Habituation Program at Bai Hokou camp, Dzanga-Ndoki NP, and all ITFC field assistants for their irreplaceable help, great humour and courage. This research was financially supported by the Leakey Foundation and the Max Planck Society.

References

Abramson, A.S. & Lisker, L. (1970). Discrimination along the voicing continuum: cross-language tests. — In: Proceedings of the 6th International Congress on Phonetic Science, Academiac, Prague, p. 569-573.

Altmann, J. (1974). Observational study of behavior: sampling methods. — Behaviour 49: 227-267.

Arnold, K. & Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. — Curr. Biol. 18: 202-203.

Arnold, K. & Zuberbühler, K. (2012). Call combinations in monkeys: compositional or idiomatic expressions? — Brain Lang. 120: 303-309.

Baugh, A.T., Akre, K.L. & Ryan, M.J. (2008). Categorical perception of a natural multivariate signal: mating call recognition in túngara frogs. — Proc. Natl. Acad. Sci. USA 105: 8985-8988.

Berwick, R.C., Okanoya, K., Beckers, G.J.L. & Bolhuis, J.J. (2011). Songs to syntax: the linguistics of birdsong. — Trends Cogn. Sci. 15: 113-121.

Blumstein, D.T. & Armitage, K.B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. — Am. Nat. 150: 179-200.

Boinski, S. (1993). Vocal coordination of troop movement among white-faced capuchin monkeys *Cebus capuchinus*. — Am. J. Primatol. 30: 85-100.

Boinski, S. & Campbell, A.F. (1996). The huh vocalization of white-faced capuchins: a spacing call disguised as a food call? — Ethology 102: 826-840.

Bouchet, H., Blois-Heulin, C., Pellier, A., Zuberbühler, K. & Lemasson, A. (2012). Acoustic variability and individual distinctiveness in the vocal repertoire of red-capped mangabeys (*Cercocebus torquatus*). — J. Comp. Psychol. 126: 45-56.

Bradbury, J.W. & Vehrencamp, S.L. (2011). Principles of animal communication. — Sinauer Associates, Sunderland, MA.

Bradley, B.J., Robbins, M.M., Williamson, E.A., Steklis, H.D., Steklis, N.G., Eckhardt, N., Boesch, C. & Vigilant, L. (2005). Mountain gorilla tug-of-war: silverbacks have limited control over reproduction in multimale groups. — Proc. Natl. Acad. Sci. USA 102: 9418-9423.

Cairns, S.J. & Schwager, S.J. (1987). A comparison of association indices. — Anim. Behav. 35: 1454-1469.

Candiotti, A., Zuberbühler, K. & Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. — Anim. Cogn. 15: 327-339.

Charlton, B.D., Ellis, W.A.H., Larkin, R. & Fitch, W.T. (2012). Perception of size-related formant information in male koalas (*Phascolarctos cinereus*). — Anim. Cogn. 15: 999-1006.

Cheney, D.L., Seyfarth, R.M. & Silk, J.B. (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. — Anim. Behav. 50: 249-257.

Clarke, E., Reichard, U.H. & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. — PLoS ONE 1: e73.

Clay, Z. & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. — PLoS ONE 6: e18786.

Cleveland, J. & Snowdon, C.T. (1982). The complex vocal repertoire of the adult cotton-top tamarin (*Saguinus oedipus oedipus*). — Ethology 58: 231-270.

Crockford, C. & Boesch, C. (2005). Call combinations in wild chimpanzees. — Behaviour 142: 397-421.

Crockford, C., Herbiger, I., Vigilant, L. & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? — Ethology 110: 221-243.

Crockford, C., Wittig, R.M., Mundry, R. & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. — Curr. Biol. 22: 142-146.

Doran, D. & McNeilage, A. (1998). Gorilla ecology and behaviour. — Evol. Anthropol. 6: 120-131.

Eens, M., Pinxten, R. & Verheyen, R.F. (1989). Temporal and sequential organisation of song bouts in the starling. — Ardea 77: 75-86.

Elowson, A.M. & Snowdon, C.T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. — Anim. Behav. 47: 1267-1277.

Fagen, R.M. & Goldman, R.N. (1977). Behavioural catalogue analysis methods. — Anim. Behav. 25: 261-274.

Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. — Anim. Behav. 55: 799-807.

Fischer, J., Hammerschmidt, K. & Todt, D. (1995). Factors affecting acoustic variation in Barbary-macaque (*Macaca sylvanus*) disturbance calls. — Ethology 101: 51-66.

Fischer, J., Noser, R. & Hammerschmidt, K. (2013). Bioacoustic field research: a primer to acoustic analyses and playback experiments with primates. — Am. J. Primatol. 75: 643-663.

Fossey, D. (1972). Vocalizations of the mountain gorilla (*Gorilla gorilla beringei*). — Anim. Behav. 20: 36-53.

Fraley, C., Raftery, A., Murphy, T.B. & Scrucca, L. (2012). MClust version 4 for R: normal mixture modelling for model-based clustering, classification, and density estimation. — Available online at http://cran.r-project.org/web/packages/mclust/index.html.

Franz, M. & Goller, F. (2002). Respiratory units of motor production and song imitation in the zebra finch. — J. Neurobiol. 51: 129-141.

Freeberg, T.M., Dunbar, R.I.M. & Ord, T.J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. — Phil. Trans. Roy. Soc. B 367: 1785-1801.

Furui, S. (1986). On the role of spectral transition for speech perception. — J. Acoust. Soc. Am. 80: 1016-1025.

Gruber, T. & Zuberbühler, K. (2013). Vocal recruitment for joint travel in wilde chimpanzees. — PLoS ONE 8: e76073.

Gustison, M.L., le Roux, A. & Bergman, T.J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. — Phil. Trans. Roy. Soc. B 367: 1847-1859.

Hammerschmidt, K. (1990). Individuelle Lautmuster bei Berberaffen (*Macaca sylvanus*): Ein Ansatz zum Verständnis ihrer vokalen Kommunikation. — Dissertation, Freie Universität Berlin, Berlin.

Hammerschmidt, K. & Fischer, J. (1998). The vocal repertoire of Barbary macaques: a quantitative analysis of a graded signal system. — Ethology 104: 203-216.

Harcourt, A.H. & Stewart, K.J. (1996). Function and meaning of wild gorilla 'close' calls 2. Correlations with rank and relatedness. — Behaviour 133: 827-845.

Harcourt, A.H. & Stewart, K.J. (2001). Vocal relationships of wild mountain gorillas. — In: Mountain gorillas. Three decades of research at Karisoke (Robbins, M.M., Sicotte, P. & Stewart, K.J., eds). Cambridge University Press, London, p. 241-262.

Harcourt, A.H., Stewart, K.J. & Harcourt, D.E. (1986). Vocalizations and social relationships of wild gorillas: a preliminary analysis. — In: Current perspectives in primate social dynamics (Taub, D.M. & King, F.A., eds). Van Nostrand Reinhold, New York, NY, p. 346-356.

Harcourt, A.H., Stewart, K.J. & Hauser, M. (1993). Functions of wild gorilla 'close calls'. I. Repertoire, context, and interspecific comparison. — Behaviour 124: 89-122.

Hohmann, G. (1989). Comparative study of vocal communication in two Asian leaf monkeys, *Presbytis johnii* and *Presbytis entellus*. — Folia Primatol. 52: 27-57.

Janik, V.M., Sayigh, L.S. & Wells, R.S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. — Proc. Natl. Acad. Sci. USA 103: 8293-8297.

Jansen, D.A.W.A.M., Cant, M.A. & Manser, M.B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls. — BMC Biol. 10: 97.

Ji, A., Johnson, M.T., Walsh, E.J., McGee, J. & Armstrong, D.L. (2013). Discrimination of individual tigers (*Panthera tigris*) from long distance roars. — J. Acoust. Soc. Am. 133: 1762-1769.

Keenan, S., Lemasson, A. & Klaus Zuberbühler, K. (2013). Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. — Anim. Behav. 85: 109-118.

Kuhl, P.K. & Miller, J.D. (1975). Speech perception by the chinchilla: voiced-voiceless distinction in alveolarplosive consonants. — Science 190: 69-72.

Lemasson, A. & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). — J. Acoust. Soc. Am. 129: 3341-3352.

Maciej, P., Fischer, J. & Hammerschmidt, K. (2011). Transmission characteristics of primate vocalizations: implications for acoustic analyses. — PLoS ONE 6: e23015.

Manser, M.B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. — Proc. Roy. Soc. Lond. B: Biol. Sci. 268: 2315-2324.

Marler, P. (2000). Origins of music and speech: insights from animals. — In: The origins of music (Wallin, N.L., Merker, B. & Brown, S., eds). MIT Press, Cambridge, MA, p. 31-48.

Masi, S., Cipolletta, C. & Robbins, M.M. (2009). Western lowland gorillas (*Gorilla gorilla gorilla*) change their activity patterns in response to frugivory. — Am. J. Primatol. 71: 91-100.

May, B., Moody, D.B. & Stebbins, W.C. (1989). Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. — J. Acoust. Soc. Am. 85: 837-847.

McComb, K. & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. — Biol. Lett. 1: 381-385.

Mitani, J.C., Gros-Louis, J. & Macedonia, J.M. (1996). Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. — Int. J. Primatol. 17: 569-583.

Mitani, J.C. & Marler, P. (1989). A phonological analysis of male gibbon singing behavior. — Behaviour 109: 20-45.

Morton, E.S. (1975). Ecological sources of selection on avian sounds. — Am. Nat. 109: 17-34.

Nelson, D.A. & Marler, P. (1989). Categorical perception of a natural stimulus continuum: birdsong. — Science 244: 976-978.

Nkurunungi, J.B., Ganas, J., Robbins, M.M. & Stanford, C.B. (2005). A comparison of two mountain gorilla habitats in Bwindi Impenetrable National Park, Uganda. — Afr. J. Ecol. 42: 289-297.

Ouattara, K., Lemasson, A. & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. — Proc. Natl. Acad. Sci. USA 106: 22026-22031.

Owren, M.J., Amoss, R.T. & Rendall, D. (2011). Two organizing principles of vocal production: implications for nonhuman and human primates. — Am. J. Primatol. 73: 530-544.

Pfefferle, D. & Fischer, J. (2003). Sounds and size — identification of variables that reflect body size in hamadryas baboons. — Folia Primatol. 74: 212.

R Core Development Team (2011). R: a language and environment for statistical computing. — R Foundation for Statistical Computing, Vienna.

Revelle, W. (2011). Psych: procedures for personality and psychological research. — R Foundation for Statistical Computing, Vienna.

Robbins, M.M. (2008). Feeding competition and agonistic relationships among Bwindi *Gorilla beringei*. — Int. J. Primatol. 29: 999-1018.

Robbins, M.M. (2010). Gorillas: diversity in ecology and behaviour. — In: Primates in perspective (Campbell, C.J., Fuentes, A., MacKinnon, K.C., Bearder, S. & Stumpf, R., eds). Oxford University Press, Oxford, p. 326-339.

Robinson, J.G. (1984). Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus olivaceus*. — Behaviour 90: 46-79.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. — Comput. Appl. Math. 20: 53-65.

Salmi, R., Hammerschmidt, K. & Doran-Sheehy, D.M. (2013). Western gorilla vocal repertoire and contextual use of vocalizations. — Ethology 119: 831-847.

Schaller, G.B. (1963). The mountain gorilla. — University of Chicago Press, Chicago, IL.

Schel, A.M., Machanda, Z., Townsend, S.W., Zuberbühler, K. & Slocombe, K.E. (2013). Chimpanzee food calls are directed at specific individuals. — Anim. Behav. 86: 955-965.

Schwarz, G.E. (1978). Estimating the dimension of a model. — Ann. Stat. 6: 461-464.

Seyfarth, R.M., Cheney, D.L., Harcourt, A.H. & Stewart, K.J. (1994). The acoustic features of gorilla double grunts and their relation to behaviour. — Am. J. Primatol. 33: 31-50.

Shapiro, A.D., Tyack, P.L. & Seneff, S. (2011). Comparing call-based versus subunit-based methods for categorizing Norwegian killer whale, *Orcinus orca*, vocalizations. — Anim. Behav. 81: 377-386.

Sicotte, P. (1994). Effect of male competition on male–female relationships in bi-male groups of mountain gorillas. — Ethology 97: 47-64.

Silk, J.B., Kaldor, E. & Boyd, R. (2000). Cheap talk when interests conflict. — Anim. Behav. 59: 423-432.

Slocombe, K.E. & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. — Curr. Biol. 15: 1779-1784.

Slocombe, K.E., Townsend, S.W. & Zuberbühler, K. (2009). Wild chimpanzees (*Pan troglodytes schweinfurthii*) distinguish between different scream types: evidence from a playback study. — Anim. Cogn. 12: 441-449.

Snowdon, C.T. & Pola, Y.V. (1978). Interspecific and intraspecific responses to synthesized pygmy marmoset vocalizations. — Anim. Behav. 26: 192-206.

Snowdon, C.T. (1997). Affiliative processes and vocal development. — Ann. N.Y. Acad. Sci. 807: 340-351.

Sokal, R.R. & Rohlf, F.J. (1995). The principles and practice of statistics in biological research. — W.H. Freeman, New York, NY.

Stebbins, W.C. & Sommers, M.S. (1992). Evolution, perception, and the comparative method. — In: The evolutionary biology of hearing (Webster, D.B., Fay, R.R. & Popper, A.N., eds). Springer, New York, NY, p. 211-227.

Stewart, K.J. & Harcourt, A.H. (1994). Gorillas' vocalizations during rest periods: signals of impending departure? — Behaviour 130: 29-40.

Sugiura, H. (2007). Effects of proximity and behavioral context on acoustic variation in the coo calls of Japanese macaques. — Am. J. Primatol. 69: 1412-1424.

Suzuki, T.N. (2014). Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. — Anim. Behav. 87: 59-65.

Tallet, C., Linhart, P., Policht, R., Hammerschmidt, K., Šimeček, P., Kratinova, P. & Špinka, M. (2013). Encoding of situations in the vocal repertoire of piglets (*Sus scrofa*): a comparison of discrete and graded classifications. — PloS ONE 8: e71841.

ten Cate, C. & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. — Phil. Trans. Roy. Soc. B 367: 1984-1994.

Waser, P.M. & Brown, C.H. (1986). Habitat acoustics and primate communication. — Am. J. Primatol. 10: 135-154.

Watts, D.P. (1991). Mountain gorilla reproduction and sexual behavior. — Am. J. Primatol. 24: 211-225.

Whitham, J.C., Gerald, M.S. & Maestripieri, D. (2007). Intended receivers and functional significance of grunt and girney vocalizations in free-ranging female rhesus macaques. — Ethology 113: 862-874.

Wich, S.A., Krutzen, M., Lameira, A.R., Nater, A., Arora, N., Bastian, M.L., Meulman, E., Morrogh-Bernard, H.C., Atmoko, S.S.U., Pamungkas, J., Perwitasari-Farajallah, D., Hardus, M.E., van Noordwijk, M. & van Schaik, C.P. (2012). Call Cultures in Orang-Utans? — PLoS ONE 7: e36180.

Wyman, M.T., Mooring, M.S., McCowan, B., Penedom, M.C.T., Reby, D. & Hart, L.A. (2012). Acoustic cues to size and quality in the vocalizations of male North American bison, *Bison bison*. — Anim. Behav. 84: 1381-1391.

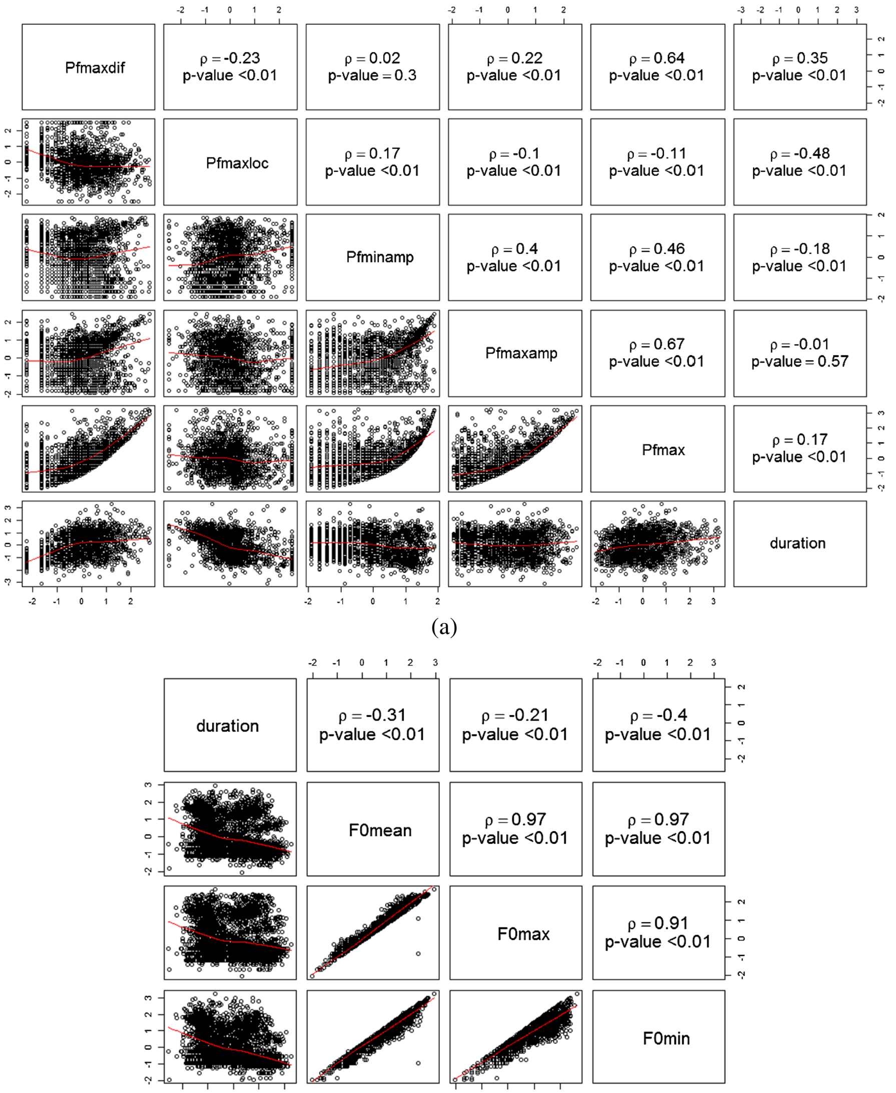
Wyttenbach, R.A., May, M.L. & Hoy, R.R. (1996). Categorical perception of sound frequency by crickets. — Science 273: 1542-1544.

Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. — Anim. Behav. 63: 293-299.

Supplementary material

A1. Principal Components Analysis

Because many of the acoustic parameters were clearly correlated to each other (Figure A1) we performed Principal Components Analyses (PCA) using the R package Psych (Revelle, 2011). Which of the given parameters were included in the PCA and how many Principal Components were calculated was based on repeated PCAs in which we successively reduced the number of included parameters and iteratively checked the Eigenvalues of the resulting components, the total variance explained by the Principal Components with Eigenvalues $\geqslant 1$ and the loadings of the acoustic parameters on them, aiming to maximize the cumulative variance explained by the derived Principal Components (following Horn, 1965). For the atonal units, the initial PCA including all measured parameters resulted in duration and Pfmaxloc almost exclusively loading on two separate components. Therefore, the PCA was rerun excluding duration and Pfmaxloc. Eigenvalues of the resulting Principal Components suggested using two Principal Components, which explained 82% of the total variance. However, after inspection of loadings, we decided to use three Principal Components as these explained 96% of the total variance (Table A1). For the tonal units, an initial PCA including all measured parameters resulted in duration almost exclusively loading on



(a)

(b)

Figure A1. Results of Spearman correlations between acoustic parameters measured for (a) atonal and (b) tonal units. (a) Correlations between acoustic parameters measured for atonal units. (b) Correlations between acoustic parameters measured for tonal units.

Table A1.
Results of the Principal Components Analysis run for atonal units.

	PC1	PC2	PC3
Pfminamp	0.03	0.2	0.98
Pfmaxamp	0.13	0.96	0.19
Pfmax	0.70	0.55	0.35
Pfmaxdif	0.98	0.06	−0.05
Eigenvalue	2.23	1.05	0.56
Cumulative variance	0.37	0.68	0.96

one component. Therefore, the PCA was rerun excluding duration. Eigenvalues of the resulting Principal Components suggested using one Principal Component, which explained 98% of the variance (Table A2).

A2. Cluster analysis

We performed cluster analyses using the R package MClust (Fraley et al., 2012) which provides functions to allow for model-based clustering. Cluster solutions were calculated through employing a maximum likelihood method that selects assignments of units to a given cluster, producing cluster centres and shapes that best fit the observed data. We calculated cluster solutions to group units into 1 to 25 clusters, and for each number of clusters we ran 10 models with different parameterizations of cluster shapes, reaching a total of 250 cluster solutions for both tonal and atonal units (for details on the parameterization of different cluster shapes see Fraley et al., 2009).

We determined the goodness of fit of the different cluster solutions using the Bayesian Information criterion (BIC), which is a criterion for model

Table A2.
Results of the Principal Components Analysis run for tonal units.

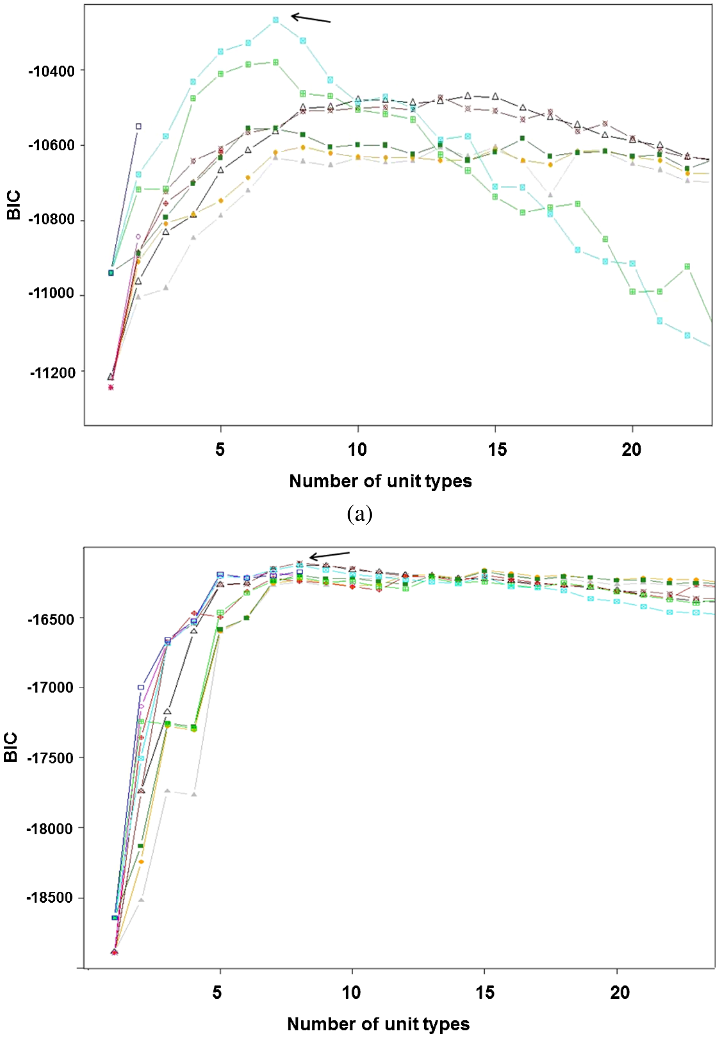
	PC1
F0max	0.99
F0min	0.98
F0mean	1.00
Eigenvalue	2.94
Cumulative variance	0.98

selection among a set of models (Schwarz, 1978), whereby an absolute difference in BIC between two models (ΔBIC) of >10 is indicative of a substantial difference between the explanatory value of the two models (Raftery, 1995). We determined the model with the best parameterizations among the 10 models with the same number of cluster and among these 25 models we determined the cluster solution with the overall best fit number of clusters. Results for atonal units suggested the cluster solution with 7 clusters to fit the data best (best model compared to the second best model: $\Delta\text{BIC} = 55.03$, Figure A2). For tonal units, results suggested the cluster solution with 8 clusters to fit the observed data best ($\Delta\text{BIC} = 18.54$, Figure A2).

In a second step, we measured the discreteness of cluster solutions by calculating silhouette coefficients, which quantify the cohesion within and separation between clusters. A $S(c_n) < 0.25$ indicates strong intergradation, $0.25 \leqslant S(c_n) < 0.5$ weak clustering, $0.5 \leqslant S(c_n) < 0.7$ medium and $S(c_n) \geqslant 0.7$ strong clustering (Rousseeuw, 1987). We calculated silhouette coefficients for the 9 cluster solutions with the number of clusters ranging from 2 to 10. We restricted this analysis to these 8 cluster solutions since the BICs indicated no improvement of model fit for cluster solutions with more than 8 clusters. For atonal units silhouette coefficients were consistently <0.2 in all cluster solutions indicating that the identified clusters were highly intergraded (Figures A3 and A4). Hence, we categorized all atonal units into one unit type. For tonal units a silhouette coefficient >0.3 indicated a weak but consistent clustering across cluster solutions, with four clusters representing the most discrete cluster solution (Figures A3 and A4).

We further validated the results of the cluster analyses by testing if clusters emerged through random regrouping of units or instead constituted subclusters of larger clusters (of solutions with fewer clusters). Therefor we investigated how units assigned to the same unit type in a given cluster solution were distributed across unit types in other cluster solutions with fewer clusters. Specifically, we determined the proportion of units belonging to the same unit type in a solution with more unit types that also belonged to the same unit type in a solution with fewer unit types.

To test whether the proportion of units remaining in the same respective unit type when the number of clusters increased was larger than expected by chance we conducted permutation tests (per pair of cluster solutions). For this we randomized the assignments of units to clusters (i.e., assignments to clusters from a solution with n clusters were randomly distributed across



(a)

(b)

Figure A2. Goodness of fit (using BIC) of different cluster solutions for atonal and tonal units as a function of the number of unit types (clusters). Differently coloured lines indicate different parameterizations of cluster solutions, varying in cluster shape. The best cluster solution is indicated by an arrow. (a) Atonal units; (b) tonal units.

the clusters of this solution). Then we determined the proportion of pairs of units that were assigned to the same respective cluster in this (randomized) solution with n clusters and also in another solution with *n*-*k* clusters (with

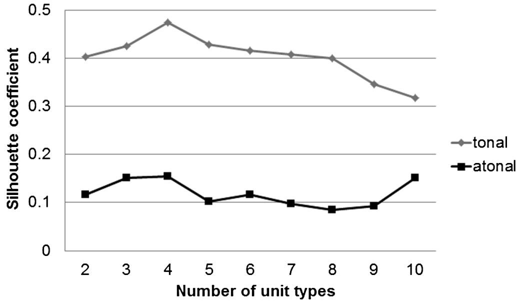


Figure A3. Silhouette coefficients of the different cluster solutions for tonal and atonal units as a function of the number of unit types (clusters). A silhouette coefficient < 0.25 indicates strong intergradation, silhouette coefficients between 0.25 to 0.5 weak clustering.

k being an integer). The *p*-value was finally determined as the proportion of permutations revealing a proportion of pairs remaining in the same respective cluster (i.e., unit type) when reducing the number of unit types being

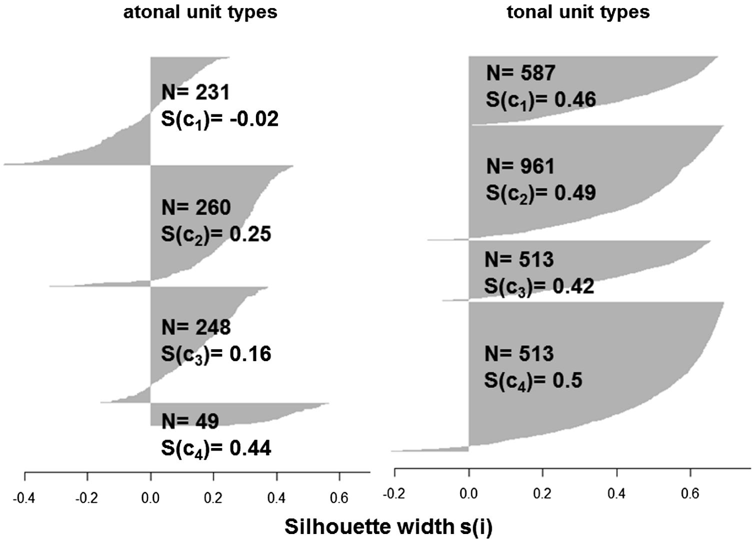


Figure A4. Silhouette plots illustrating the most discrete cluster solutions with four unit types (clusters) for atonal units, with an overall silhouette coefficient of 0.15 (left) and with four unit types for tonal units, with an overall silhouette coefficient of 0.48 (right). The plot displays the silhouette widths for each unit as horizontal bars, ordered by clusters and by silhouette value within cluster. The silhouette width is a measure of how clearly a unit belongs to its assigned cluster. Compact clusters consist of units with high silhouette widths. N indicates the number of units allocated to each cluster. $S(c_n)$ is the silhouette coefficient of each cluster.

at least as large as for the original data (e.g., Good, 2005). For the atonal units cluster analyses differing in the assumed number of clusters did not lead to arbitrarily different arrangements of units. Additional clusters identified in cluster solutions with more clusters largely constituted sub-clusters within clusters identified by cluster solutions with fewer clusters. In fact, an average of 71% of units belonging to the same respective cluster in a given cluster solution were also assigned to the same cluster in another cluster solution with fewer clusters, which was significantly higher than expected by chance (permutation tests, all $p = 0.001$). Also for the tonal units cluster solutions differing in the assumed number of clusters did not lead to arbitrarily different arrangements of units. Additional unit types identified in cluster solutions with more clusters largely constituted sub-clusters of clusters identified by cluster solutions with fewer clusters. In fact, an average of 92% of units belonging to the same respective cluster in a given cluster solution were also assigned to the same cluster in another cluster solution with fewer unit types (permutation tests, all $p = 0.001$, Figure A5).

A3. Results based on the classification of units into 15 unit types

A3.1. Repertoire of unit types

Individuals showed a median repertoire size of 13 unit types (mountain gorilla group: 8 to 14, western gorilla group: 11 to 15). None of the 15 unit types was exclusively found in only one individual, or in only one of the two groups. Nevertheless, we found a male specific use of the atonal unit type 7 (a7) in both groups. Furthermore, tonal unit type 4 (t4) was highly specific (but not unique) to the one western gorilla blackback (Table A3). Our sampling was sufficient to compile an almost complete repertoire of unit types since the average sample coverage for the individuals of the mountain gorilla group was 0.99 (0.96 to 1, $N = 10$ individuals) and 0.99 for the individuals of the western gorilla group (0.98 to 1, $N = 5$ individuals). Differences between individual repertoire sizes were not due to differences in sample sizes since we found no correlation between the number of recorded unit types and the number of analysed units (mountain gorilla group: Spearman correlation, $\rho = 0.41$, $N = 10$ individuals, $p = 0.24$; western gorilla group: Spearman correlation, $\rho = 0.051$, $N = 5$ individuals, $p = 1$).

Two given individuals shared on average ca. 85% of their repertoire of unit types, as indicated by a high average Dice index of repertoire overlap of 0.88 across all dyads (0.63 to 1, $N = 105$ dyads). The similarity between

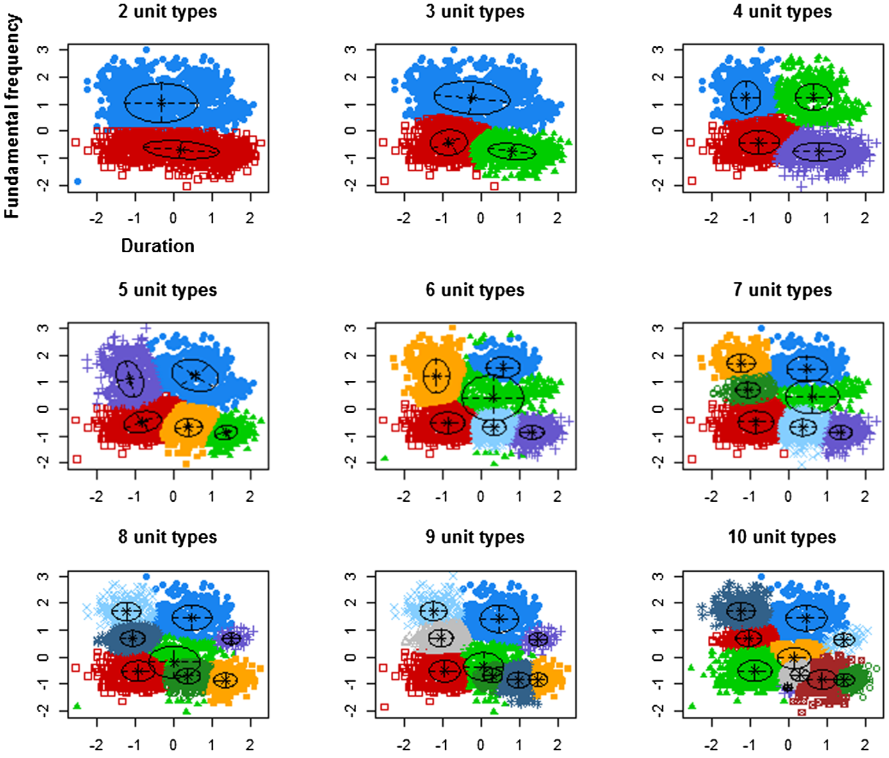


Figure A5. Best cluster solutions for 2 to 10 unit types (clusters) of tonal units. Indicated are the values of the principal component for acoustic parameters describing the fundamental frequency on the y-axis and the transformed duration on the x-axis. Different colours and symbols indicate different clusters. For each cluster the cluster mean is marked and ellipses indicate corresponding covariance. Note that an increase in the number of clusters usually largely lead to a split of a certain cluster rather than a more fundamental change in the arrangement of units into clusters.

repertoires of individuals was larger when they belonged to the same age–sex class (blackbacks of the mountain gorilla group: 0.98 (0.96 to 1), $N = 4$; adult females of the mountain gorilla group: 0.86 (0.67 to 1), $N = 5$) compared to when they belonged to different age–sex classes (0.87 (0.73 to 0.96); permutation test, $p = 0.001$). The similarity between repertoires was not obviously related to whether females belonged to the same group or not (permutation test, $N = 5$ adult females of the mountain gorilla group, 3 adult females of the western gorilla group, $p = 0.45$).

Table A3.
Overview of the repertoire of unit types and combinations found in close calls of the individuals of the mountain gorilla and the western gorilla group.

	Mountain gorilla											Western gorilla						Total
	AF					BB				SB	Total mountain gorilla	AF			BB	SB	Total western gorilla	
	BY	KK	MG	ST	TN	BZ	FZ	MR	SK	RC		BE	MAL	MPB	KU	MK		
a1	14	2	2	2	0	23	7	4	1	7	62	13	9	1	3	6	32	94
a2	11	0	2	4	0	4	5	1	2	49	78	8	1	5	4	8	26	104
a3	12	3	6	5	0	17	3	0	2	16	64	1	4	4	2	1	12	76
a4	15	3	3	5	0	23	25	11	1	22	108	16	21	13	17	10	77	185
a5	10	1	4	0	1	27	14	15	9	37	118	8	2	0	0	7	17	135
a6	12	5	1	5	1	17	3	3	2	18	67	11	16	17	9	10	63	130
a7	0	0	0	0	0	15	2	5	1	6	29	0	0	0	8	10	18	47
t1	30	18	61	21	10	92	58	26	34	0	350	5	0	1	36	21	63	413
t2	30	56	62	7	2	43	64	36	41	12	353	27	203	79	71	42	422	775
t3	9	18	45	26	6	1	1	4	1	34	145	9	1	0	51	10	71	216
t4	0	0	0	1	0	0	0	0	0	0	1	1	0	0	55	1	57	58
t5	6	52	49	31	0	6	42	2	4	6	198	66	142	50	0	134	392	590
t6	10	4	12	16	19	79	40	41	17	1	239	3	0	1	32	5	41	280
t7	18	92	51	33	10	88	100	62	28	25	507	16	16	9	3	55	99	606
t8	1	15	17	4	4	3	4	4	5	5	62	2	3	4	163	14	186	248
No. of calls	90	156	154	90	27	212	171	99	77	124	1200	133	266	121	188	222	930	2130
No. of units	178	269	315	160	53	438	368	214	148	238	2381	186	418	184	454	334	1576	3957
No. of unit types	13	12	13	13	8	14	14	13	14	13	15	14	11	11	13	15	15	15
No. of combinations	57	42	54	31	13	67	58	43	30	40	200	44	29	33	88	57	175	305
Sampling coverage unit types	0.99	1	1	0.99	0.96	1	1	1	0.97	1	0.99	0.99	1	0.98	1	0.99	0.99	0.99
Sampling coverage combinations	0.59	0.83	0.81	0.83	0.67	0.83	0.78	0.7	0.71	0.87	0.76	0.83	0.93	0.81	0.68	0.83	0.82	0.78

AF: adult female, BB: blackback, SB: silverback; BY, KK, MG, ST, TN, BZ, FZ, MR, SK, RC, BE, MAL, MPB, KU, MK denote individual gorillas.

A3.2. Repertoire of combinations

Overall, the 15 unit types were assembled into 305 different combinations (Table A3). We recorded 200 combinations for the individuals of the mountain gorilla group with an average of 44 combinations per individual (13 to 67, $N = 10$ individuals). In the individuals of the western gorilla group we found 175 combinations with an average of 50 combinations per individual (29–88, $N = 5$ individuals). We were close to having sampled the complete set of combinations since the average sample coverage was 0.76 for the individuals of the mountain gorilla group (0.59–0.87, $N = 10$ individuals) and 0.82 for the individuals of the western gorilla group (0.68–0.93, $N = 5$ individuals). Differences between individual repertoire sizes in the mountain gorilla group were likely due to differences in sample size, since the number of recorded combinations was positively correlated with the number of calls analysed for the mountain gorilla individuals (Spearman correlation, $\rho = 0.79$, $N = 10$ individuals, $p = 0.0075$), but not for the individuals of the western gorilla group (Spearman correlation, $\rho = -0.1$, $N = 5$ individuals, $p = 0.95$). Individuals shared on average ca. 25% of their repertoire of combinations as indicated by an average Dice coefficient of repertoire overlap of 0.26 (0–0.52, $N = 105$ dyads). Similarity between repertoires of two given females was larger when they were of the same group (adult females of the mountain gorilla group: 0.3 (0.14–0.44), $N = 5$; adult females of the western gorilla group: 0.4 (0.36–0.42), $N = 3$), as compared to females that belonged to different groups (0.21 (0–0.46); permutation test, $p = 0.007$). The similarity between repertoires was not obviously related to whether individuals belonged to the same age–sex class or not (permutation test, $N = 5$ adult females of the mountain gorilla group, 4 blackbacks of the mountain gorilla group, $p = 0.31$).

A3.3. Combinatorial association between unit types

The presence of a given unit type within a call was only weakly associated with the presence of other unit types within the same call as indicated by a low average Dice index of combinatorial association of 0.05 across individuals (0.03–0.09, $N = 15$ individuals). This shows that unit types were flexibly combined into calls. However, association strengths differed between unit types. While across individuals 42–77% of possible pairs of unit types were never observed in the same call (Dice index = 0), association strengths between unit types that were observed in the same call, ranged from an average minimum Dice index of 0.03 (0.01–0.09, $N = 15$ individuals) to an

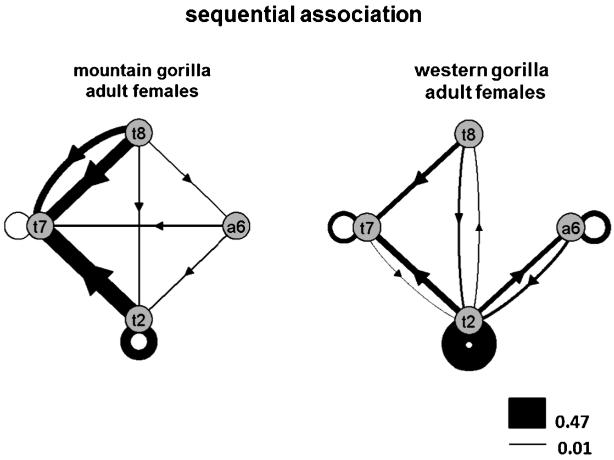


Figure A6. Sequential associations between unit types in the adult females of both gorilla groups. Lines indicate the average Dice index of sequential association as a measure of how frequently any two unit types followed one another in immediate succession. Average values were calculated from the individuals' Dice indices for a given pair of unit types. The legend depicts the thickness of the lines corresponding to the minimum and maximum average Dice index. Included are only unit types shared by all adult females.

average maximum of 0.45 (0.28–0.63, $N = 15$ individuals). Overall association strength did not differ between groups (exact Mann–Whitney U-test, $U = 12$, $N = 5$ adult females of the mountain gorilla group, 3 adult females of the western gorilla group, $p = 0.25$) or age–sex classes ($U = 12$, $N = 4$ blackbacks of the mountain gorilla group, 5 adult females of the mountain gorilla group, $p = 0.73$).

A3.4. Sequential association between unit types

The sequential arrangement of unit types within calls was non-random as indicated by an average contingency coefficient of 0.79 across the individual transition matrices (0.65–0.88, $N = 15$ individuals). Some unit types followed each other in immediate succession more frequently than others. While across individuals 0% to 43% of the possible transitions between unit types were never observed (Dice index of sequential association = 0), for observed transitions between unit types, association strengths ranged from an average minimum Dice index of 0.05 (0.02–0.1, $N = 15$ individuals) to an average maximum of 0.6 (0.43–1, $N = 15$ individuals). Individuals arranged the same unit types into similar sequences as we observed some inter-individual consistency among the transition probabilities between unit types. In fact, there was a significant positive correlation for 59% of the 105 comparisons between transition matrices of two individuals with an average correlation coefficient of 0.37 (0.22–0.6; Spearman correlation, $p \leqslant 0.05$). Correlation coefficients of individual transition matrices where larger when females belonged to the same group (adult females of the mountain gorilla group: 0.36 (0.16–0.6), $N = 5$; adult females of the western gorilla group: 0.36 (0.3–0.45), $N = 3$) than when females belonged to different groups (0.15 (-0.12 to 0.35); permutation test, $p = 0.018$; Figure A6). Correlation coefficients also tended to be larger when individuals belonged to the same age–sex class (blackbacks of the mountain gorilla group: 0.5 (0.45–0.57), $N = 4$; adult females of the mountain gorilla group: 0.36 (0.17–0.6), $N = 5$) than when they belonged to different age–sex classes (0.35 (0.1–0.6); permutation test, $p = 0.09$).

References

Horn, J.L. (1965). A rationale and test for the numbp-valueer of factors in factor analysis. — Psychometrika 30: 179-185.

Fraley, C., Raftery, A. & Scrucca, L. (2012). MClust: Normal mixture modelling for model-based clustering, classification, and density estimation. — R Foundation for Statistical Computing, Vienna.

Raftery, A.E. (1995). Bayesian model selection and social research. — Sociol. Methodol. 25: 111-163.

Revelle, W. (2011). Psych: Procedures for Personality and Psychological Research. — R Foundation for Statistical Computing, Vienna.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. — Comput. Appl. Math. 20: 53-65.

Schwarz, G.E. (1978). Estimating the dimension of a model. — Ann. Stat. 6: 461-464.