

# Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax?

Daniela Hedwig · Roger Mundry ·  
Martha M. Robbins · Christophe Boesch

Received: 12 May 2014 / Revised: 22 September 2014 / Accepted: 25 September 2014 / Published online: 14 October 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The core of the generative power of human languages lies in our ability to combine acoustic units under specific rules into structurally complex and semantically rich utterances. While various animal species concatenate acoustic units into structurally elaborate vocal sequences, such compound calls do not appear to be compositional as their information content cannot be derived from the information content of each of its components. As such, animal compound calls are said to constitute a form of phonological syntax, as in the construction of words in human language, whereas evidence for rudimentary forms of lexical syntax, analogous to the construction of sentences out of words, is scarce. In a previous study, we demonstrated that the repertoire of close-distance calls of mountain and western gorillas consists of acoustic units that are either used singularly or non-randomly combined. Here, we investigate whether this syntactic variation provides indications for lexical or phonological syntax. Specifically, we examined the differences between the potential information content of compound calls and their components used singularly through investigating the contexts in which they are used. We found that the gorillas

emitted compound calls in contexts similar to some but not all components, but also in a context rarely found for any of their components. As such, the investigated compound calls did not appear to be compositional as their information content cannot be derived from the information content of each of their components. Our results suggest that combining acoustic units into compound vocalizations by gorillas constitutes a form of phonological syntax, which may enable them to increase the number of messages that can be transmitted by an otherwise small repertoire of acoustic units.

**Keywords** Syntax · Mountain gorilla · Western gorilla · Close-distance vocalizations

## Introduction

One defining criterion of human language is its vast expressive power as opposed to the semantically limited vocal repertoires of non-human animal species (Chomsky 1965; Jackendoff 1999; Nowak et al. 2000). The core of the generative power of human languages lies in our ability to flexibly combine acoustic units under specific rules into structurally highly complex and semantically rich utterances. This is achieved through two main syntactic principles. First, via phonological syntax, we construct meaningful words out of phonemes that are not themselves meaningful. Second, via lexical syntax, we generate *compositional* messages, i.e., phrases and sentences, to which each constituent word contributes its own independent meaning (e.g., Marler 1977; Jackendoff 1999; Hurford 2011). Finding the evolutionary origins of the syntactic capacity of humans is one of the challenges in order to further our understanding of the origins of human

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10071-014-0812-6) contains supplementary material, which is available to authorized users.

---

D. Hedwig (✉) · R. Mundry · M. M. Robbins · C. Boesch  
Max Planck Institute for Evolutionary Anthropology,  
Department of Primatology, Deutscher Platz 6, 04103 Leipzig,  
Germany  
e-mail: hedwig@eva.mpg.de

R. Mundry  
Max Planck Institute for Evolutionary Anthropology,  
Department of Developmental and Comparative Psychology,  
Deutscher Platz 6, 04103 Leipzig, Germany

language, which up to date remains highly disputed (Hauser et al. 2002).

Syntactic variation appears to be an evolutionarily deeply rooted principle of how information is conveyed, with compound vocalizations being notably widespread among animal species ranging from birds to mammals. Marler (1977) recognized two types of syntax, analogous to phonological and lexical syntax in human language. Phonological syntax, defined as the concatenation of sounds without independent information content and which are not used singularly, or meaningful sounds that lose their original information content when combined (Marler 1977), appears to be widespread (e.g., birds: Marler and Mitani 1988; *Hyllobates agilis*: Mitani and Marler 1989; *Pteronotus parnellii*: Kanwal et al. 1994; *Procapra capensis*: Kershenbaum et al. 2012). In contrast, evidence for lexical syntax, through which the meaning of compound vocalizations arises from the independent meaning of each of its parts (Marler 1977), seems to be confined to primate species. These compound vocalizations occur in situations that comprise the contexts of each of their units used singularly. Therefore, these vocalizations appear to be compositional messages since their information content seems to reflect the information content of each of their units. On the one hand, such compound calls are given in situations in which the contexts of component calls are co-occurring, in that the information contents of the component calls are combined in a purely additive fashion (e.g., pant hoot and grunt combinations in *Pan troglodytes*: Crockford and Boesch 2005; close call combinations in *Cercopithecus diana*: Candiotti et al. 2012). On the other hand, lexical compound calls appear to act as contextual modifiers as they are produced in intermediary, intensified, or alleviated contexts compared to the contexts they share with their constituent calls. For instance, call combinations occur only in the contexts in which the range of contexts of the constituent calls overlaps (*Cebus olivaceus*: Robinson 1984); alarm call combinations are given in alarm situations that require vigilance but with reduced fear (*Saguinus oedipus*: Cleveland and Snowdon 1982; *Cercopithecus diana*: Zuberbühler 2002; Ouattara et al. 2009); and food call combinations appear to be given in food contexts with increased arousal state (*Pan troglodytes*: Crockford and Boesch 2005, but see Schassburger 1993 for threat call combinations in wolves functioning as intensified threat signals).

To date, research on syntactic variation in primate vocalizations was mainly focused on monkeys and specifically their predator alarm call systems (e.g., *Cercopithecus diana*: Zuberbühler 2002; *Cercopithecus campbelli*: Ouattara et al. 2009). Remarkably, little research has considered syntactic structures in vocalizations of great apes (but see Crockford and Boesch 2005; Clay and Zuberbühler 2011; Lameira et al. 2013), which are, because of their close phylogenetic proximity to humans, particularly

relevant to our understanding of the evolution of the syntactic capacity of humans. Furthermore, given the limited acoustic and contextual variability of alarm calls, the importance of predation pressure as driving force on the selection toward syntactic communication in the human lineage is currently debated (e.g., Stephan and Zuberbühler 2008; Lameira et al. 2013). Studying the syntactic properties of primate close-distance vocalizations, which regulate many aspects of primate social behavior (e.g., Boinski 1993; Cheney et al. 1995; Boinski and Campbell 1996; Silk et al. 2000; Whitham et al. 2007), seems adequate given the inherently social nature of human language. Close-distance vocalizations are acoustically highly variable (Elowson and Snowdon 1994; Sugiura 2007; Lemasson and Hausberger 2011; Bouchet et al. 2012), which has been suggested to be beneficial in the face of variable influences of numerous social factors (e.g., Snowdon 1997; Bouchet et al. 2012). A recent study suggests that such flexibility in close range social calls can arise through combinatorial systems (e.g., Candiotti et al. 2012).

Gorillas are an excellent great ape model to study syntactic variation in close-distance vocalizations, as so-called close calls are characteristic for gorilla vocal behavior (e.g., Harcourt et al. 1993; Harcourt and Stewart 1996; Salmi et al. 2013). They form 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). Gorillas give close calls usually as part of vocal exchanges and most frequently when individuals are in close proximity, including in potentially agonistic situations when feeding close to each other (Harcourt et al. 1986, 1993). However, close calls are also emitted during affiliative interactions when individuals are resting, during grooming, and during interactions with infants, as well as in situations of potential separation when individuals are far from others and before changes in group activity (Harcourt et al. 1986, 1993; Seyfarth et al. 1994; Salmi et al. 2013). Based on this, various functions for the different gorilla close calls have been suggested, ranging from facilitating friendly interactions or appeasement over maintenance of group cohesion to mildly agonistic signals mediating feeding competition (Schaller 1963; Fossey 1972; Harcourt et al. 1993; Seyfarth et al. 1994; Stewart and Harcourt 1994; Harcourt and Stewart 1996, 2001). Formal testing of these functional hypotheses on a detailed acoustic level will be necessary for a better understanding of the gorilla close call system. Moreover, gorillas live in a wide variety of ecological conditions that should lead to variation in social and hence also vocal behavior (Robbins 2010). However, most of what we know about gorilla vocal behavior derives from the well-studied mountain gorillas, whereas our knowledge about western gorilla vocal

behavior is still limited. A recent study by Salmi et al. (2013) suggested contextual differences in the use of close calls in the two gorilla species. However, directly comparable data that allow linking acoustic to contextual variation in the two gorilla species have so far not been available.

The purpose of this study was to investigate contextual correlates of syntactic variation in close calls of mountain and western gorillas. In a previous study, Hedwig et al. (2014) have demonstrated that gorilla close calls consist of five acoustic unit types: low-pitched short atonal grunts (in the following referred to as A1) and tonal grunts (T2), longer grumbles (T4), as well as high-pitched short hums (T1) and long hums (T3). The gorillas used these unit types singularly or combined into a total of more than 150 combinations, which were produced in a flexible but non-random manner (Hedwig et al. 2014), similar to other primate species' syntactic systems (e.g., *Saguinus oedipus*: Cleveland and Snowdon 1982; *Cebus olivaceus*: Robinson 1984; *Pan troglodytes*: Crockford and Boesch 2005). The most frequent combinations were the low-pitched unit types A1, T2, and T4, either used singularly or as pairwise combinations, which were used by all age–sex classes in both the western and mountain gorilla group investigated. These calls are equivalent to what has previously been termed as single grunt, double grunts, and grumbles, which are the vocalizations most frequently emitted by gorillas (Harcourt et al. 1993; Salmi et al. 2013). Given the frequent use of these unit types and their combinations, they seem to play a particularly important role in gorilla social behavior and hence were the focus of our study.

First, we investigated whether gorilla close calls provide indications for lexical or phonological syntax. Overall, since in syntactic systems the information content of a combination should be different from the information content of unit types used alone, we expected contextual differences between unit types used alone and in combination with other unit types (e.g., Cleveland and Snowdon 1982; Robinson 1984; Crockford and Boesch 2005; Candiotti et al. 2012). If close calls represent a form of lexical syntax, we predicted them to be *compositional*, in the sense that the information content of a combination will reflect the information content of its constituent unit types used singularly. This would be indicated by an overlap between the contexts of a combination and each of its constituent unit types used singularly. Moreover, lexical syntax will be indicated by a correlation between the context and the variation in the sequential order of unit types within a call (e.g., “hunt chimpanzee” is different from “chimpanzee hunt,” Crockford and Boesch 2005). In turn, if close calls constitute a form of phonological syntax, the information content of a combination should

not reflect the information content of its constituent unit types used singularly. This should be indicated by combinations occurring in contexts different from the context of their unit types used singularly, including combinations comprising their repetition. Compositionality (the meaning of a sentence can be derived from the meanings of its parts) constitutes the hallmark feature underlying lexical syntax in human speech (e.g., Portner 2005). This principle of compositionality is violated by human language in rare cases as, for instance, the figurative meaning of an idiomatic phrase cannot be deduced from the meaning of its parts (e.g., “to beat a dead horse”). However, assuming that such complexity evolved gradually, we would expect not the same level of sophistication in gorilla close call syntax.

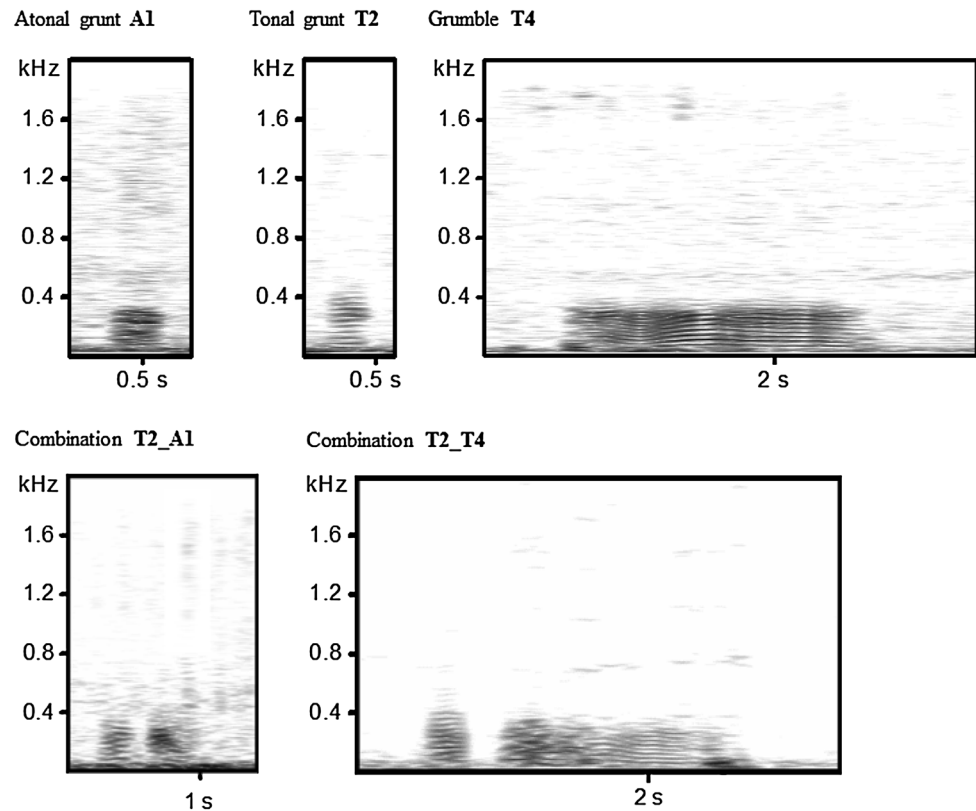
Second, we investigated possible functions of gorilla close calls as suggested in previous studies (Schaller 1963; Fossey 1972; Harcourt et al. 1993; Seyfarth et al. 1994; Stewart and Harcourt 1994; Harcourt and Stewart 1996, 2001). We predicted that calls functioning to signal a caller's location will be more likely during foraging, when individuals are on the move, and when individuals are far from other group members. Calls involved in facilitating friendly social interactions will be more likely when individuals are resting in close proximity to others, which is when affiliative social interactions are most likely, and particularly in close proximity to an infant, since close-distance calls triggered by the presence of infants have been suggested to serve as signals of benign intent in other primate species (e.g., *Papio cynocephalus ursinus*: Cheney et al. 1995; *Macaca mulatta*: Whitham et al. 2007). Moreover, calls functioning to mediate feeding competition should be given when animals are feeding in close proximity to each other. Lastly, we tested for contextual differences between the close calls of the two gorillas species (Salmi et al. 2013).

## Methods

### Study sites and subjects

Data collection on mountain gorillas focused on 10 adult individuals (one silverback, five females, and four blackbacks) of the habituated “Kyagurilo” group (Robbins 2008) in Bwindi Impenetrable National Park, Uganda. Due to the regulations of the Uganda Wildlife Authority, observations were restricted to approximately 4 h 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. For western gorillas, data were collected on 5 adult individuals (one silverback, three females, and one blackback) of the “Makumba” group

**Fig. 1** Spectrograms of representative examples of the unit types A1, T2, and T4 (all western gorilla adult female), as well as their combinations T2\_A1 (mountain gorilla adult female) and T2\_T4 (mountain gorilla blackback male). Atonal grunt A1: average duration 0.32 s (range: 0.03–2.17), average maximum peak frequency 184 Hz (74–384); tonal grunt T2: average duration 0.22 s (range: 0.03–0.63), average mean fundamental frequency 40 Hz (13–82); grumble T4: average duration 1.34 s (0.34–5.01), average mean fundamental frequency 31 Hz (11–65) (from Hedwig et al. 2014)



(Masi et al. 2009) observed at the Bai Hokou study site located in the Dzanga-Ndoki National Park (Dzanga Sangha Protected Areas), Central African Republic. Data were collected in the mornings (7:00–12:00) and/or afternoons (12:00–17:00) on 124 days over a period of 8 months from April to November 2009.

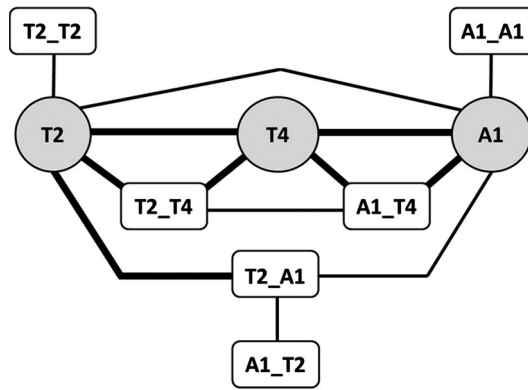
#### Data collection

We used focal animal sampling (Altmann 1974) during which we conducted continuous audio recordings of the focal animal's vocal behavior. An average of 33 h of audio recordings was collected for each adult mountain gorilla individual (range 32.6–34.6) and an average of 58 h for each adult western gorilla individual (range 57–58.5). For each vocalization, we recorded the following data describing the immediate context of call emission: (1) the activity of the focal animal, which we categorized as resting (sleeping, sitting, grooming, nursing) or foraging (preparation and ingestion of food, locomotion between feeding spots); (2) whether there was a weaned or infant (non-weaned) individual within 5 m of the caller (y/n); and (3) whether the vocalization was given spontaneously or in reply to another individual's call. A reply was defined as any vocalization given within 3 s after another individual's call (based on Seyfarth et al. 1994).

#### Statistical analysis

The analysis was based on a subset of calls for which we demonstrated the syntactic properties of gorilla close calls (Hedwig et al. 2014). Our study focused on the most commonly observed unit types, the atonal grunt A1, the tonal grunt T2 and the grumble T4, and their most common combinations since these seem to play a particularly important role in gorilla social behavior given their frequent use (Fig. 1). These combinations were observed to be given by a minimum of 9 of the 15 sampled individuals with individual call rates ranging from 0 to 11.9 calls per hour (A1: average call rate = 0.4 (0–1.7), observed in 12 individuals; T2: 0.6 (0–5.2), 9 individuals; T4: 2.7 (0–11.9), 13 individuals; A1\_A1: 1 (0–6.2), 12 individuals; T2\_T2: 0.7 (0–3.5), 9 individuals; A1\_T2: 0.2 (0–1.2), 11 individuals; T2\_A1: 0.5 (0–2.2), 13 individuals; A1\_T4: 0.9 (0–2.5), 14 individuals; T2\_T4: 1.4 (0–3.6), 14 individuals). Among these, we selected the calls for which we had all necessary contextual information and a sufficient number of observations to allow for a meaningful statistical test of our hypotheses. A total of 1321 calls were considered for analysis (see Table 1 for additional information on the number of individuals and calls used for each analysis). We ran four sets of analyses. We tested for contextual differences between (1) the unit types

A1, T2, and T4 used singularly; (2) A1, T2, and T4 used singularly and their respective combinations; (3) combinations comprising unit types A1 and T2 in reversed order; and (4) unit types A1 and T2 used singularly and in combinations comprising repetitions of either A1 or T2 (Fig. 2; Table 1).



**Fig. 2** Overview of comparisons made in this study. Different comparisons are indicated by the respective connecting lines. *Gray circles* indicate unit types used singularly and boxes depict their respective combinations. Sample sizes did not allow to conduct all possible comparisons between the illustrated combinations. *Thin lines* indicate nonsignificant comparisons, *bold lines* significant comparisons (Table 1). Comparisons were carried out using generalized linear mixed models (see text for more details)

We used generalized linear mixed models with binomial error distribution and a logit link function (GLMM; Baayen 2008) for the binary response variable coding one of the combinations to be compared as “1” and the other one as “0.” Models were fitted using the function “glmer” provided by the lme4 package (Bates et al. 2013) in the statistical software environment R, version 3.0.2 (R Core Team 2013). Into the models, we included context variables as fixed effects and “caller” as a random effect. To differentiate between competitive situations (foraging in close proximity to another individual) and socially relaxed, potentially affiliative situations (resting in close proximity to another individual), we included an interaction between “individual within 5 m” and “activity.” We expanded this by specifically including a three-way interaction between “individual within 5 m,” “activity,” and “group” to test for contextual difference in the two study groups, as suggested by Salmi et al. (2013). Lastly, we included an autocorrelation term to account for temporal interdependence between data points (Table 2). A significant autocorrelation term indicates that observations made temporally close to one another are similar (see Online Resource 1 for detailed information on how the autocorrelation term was calculated).

For all models, we tested for collinearity between the predictor variables by calculating variance inflation factors (VIF, Field 2005), using the function vif of the R package

**Table 1** Overview of models run to compare the different unit types and their combinations, the respective sample sizes, and the results of likelihood ratio tests comparing full models including all predictor

variables with reduced models comprising only the autocorrelation term and the random effect of “caller”

Comparisons	# Individuals	# Observations	Likelihood ratio test
<i>Unit types used singularly</i>			
1 T4 T2	8 (wg:4, mg:4)	414 (T4:336, T2:78)	$\chi^2 = 41.22, df = 9, p < 0.001$
2 T4 A1	9 (wg:4, mg:5)	417 (A1:48, T4:374)	$\chi^2 = 28.36, df = 9, p < 0.001$
3 A1 T2	5 (wg only)	107 (A1:33, T2:74)	$\chi^2 = 1.54, df = 5, p = 0.909$
<i>Unit types combined versus singularly</i>			
4 T2_T4 T2	8 (wg:4, mg:4)	208 (T2:130, T2_T4:78)	$\chi^2 = 22.32, df = 9, p = 0.008$
5 T2_T4 T4	13 (wg:4, mg:9)	676 (T4:431, T2_T4:245)	$\chi^2 = 48.31, df = 9, p < 0.001$
6 A1_T4 T4	13 (wg:4, mg:9)	567 (T4:431, A1_T4:136)	$\chi^2 = 30.61, df = 9, p < 0.001$
7 A1_T4 A1	10 (wg:5, mg:5)	158 (A1:104, A1_T4:54)	$\chi^2 = 17.91, df = 9, p = 0.036$
8 T2_A1 T2	8 (wg:5, mg:3)	137 (T2:58, T2_A1:79)	$\chi^2 = 18.89, df = 9, p = 0.026$
9 T2_A1 A1	10 (wg:5, mg:5)	122 (A1:68, T2_A1:54)	$\chi^2 = 14.04, df = 9, p = 0.121$
10 T2_T4 A1_T4	13 (wg:4, mg:9)	381 (T2_T4:245, A1_T4:136)	$\chi^2 = 10.24, df = 9, p = 0.332$
<i>Sequential order</i>			
11 T2_A1 A1_T2	9 (wg:4, mg:5)	85 (T2_A1:60, A1_T2:25)	$\chi^2 = 15.41, df = 9, p = 0.081$
<i>Repetitions</i>			
12 T2 T2_T2	7 (wg:5, mg:2)	172 (T2:94, T2_T2:78)	$\chi^2 = 14.24, df = 9, p = 0.114$
13 A1 A1_A1	9 (wg:5, mg:4)	144 (A1:92, A1_A1:52)	$\chi^2 = 12.07, df = 9, p = 0.209$

Sample sizes differ because we excluded subjects with very small samples for the respective unit types or combinations

wg western gorilla, mg mountain gorilla



**Table 2** Fixed and random effects included into models used to compare the contexts of the different unit types A1, T2 and T4, and their combinations

	Fixed effects	Random effects
Activity*	Categorical: foraging (0), resting (1)	Random intercepts of “caller”
Individual within 5 m*	Categorical: no (0), yes (1)	
Group +	Categorical: mountain gorilla (0), western gorilla (1)	Random slopes of significant predictor variables within “caller”
Infant within 5 m +	Categorical: no (0), yes (1)	
Vocal role +	Categorical: reply call (0), spontaneous call (1)	
Autocorrelation term	Continuous	

\* Three-way and all two-way interactions between the respective main effects activity, individual within 5 m and group

“car” (Fox and Weisberg 2011) based on a standard linear model excluding the random effect (R function “lm”). We found a maximum VIF = 2.03 across models and predictors, which is considered to be unproblematic (Field 2005).

As an overall test of the influence of our context variables on the probabilities of the combinations, we compared the full model with all predictor variables, to a null model only including the autocorrelation term and the random effects variable “caller” (Forstmeier and Schielzeth 2011) using a likelihood ratio test (Dobson 2002) and the R function “anova”. In case the full-null model comparison revealed significance and the interactions did not, we excluded them (beginning with the three-way interaction) to obtain interpretable *p* values for the respective main effects or two-way interactions.

To rule out that certain individuals were particularly influential for the results, we removed individuals one at a time from the data, ran the model again, and compared the estimated coefficients with those derived from the data set including all individuals. Across models, estimates derived from data sets with individuals removed in part differed considerably from estimates derived from models based on all data. This was particularly the case for predictors being nonsignificant in the model based on all data. From this, we concluded our nonsignificant findings need to be treated somewhat cautiously (i.e., effects could actually be stronger than our models suggest). However, we did not detect any indication that our data sets contained individuals that could destabilize the estimates of significant predictor variables.

Due to complete separation, i.e., the response being completely explained by a specific predictor (Field 2005), model estimates could sometimes not reliably be calculated (indicated by heavily inflated estimates with extremely large standard errors for some of the predictors). In order to derive interpretable estimates for the respective predictors, the problematic cases were identified and the following randomization method was applied. We reran a given model a number of times corresponding to the number of problematic cases (the number of times the response always being “0” in a specific context), whereby for each model we exchanged one of the “0” cases with a “1.”

Model estimates were then derived as the mean of the estimates derived from each of these models. Note that this approach is conservative in that it weakens the effect of the respective predictor on the response.

Following Barr et al. (2013), we aimed to include random intercepts and random slopes for each predictor variable to model random variation in the mean response and the strength of the effect of the predictors among the different individuals. However, complete separation of the response variable for the levels of the predictor variable in the majority of individuals and models made it impossible to include random slopes for all the predictor variables simultaneously. Therefore, we first ran all models with only the random intercept of “caller.” To avoid model instability, we ran models only including individuals for which at least one observation for both of the combinations to be compared was available. Subsequently, we reran the models with random slopes for the significant predictor variables. In those cases, we present the results of the model without random slopes for the nonsignificant predictors and for the significant predictors the results of the model including its random slope. This approach seems justified since neglected random slopes lead to *p* values being too small (Schielzeth and Forstmeier 2009; Barr et al. 2013), so they should not be an issue for nonsignificant predictors.

## Results

### Unit types used singularly

Overall, we found contextual differences in the probability of occurrence of T4 compared to both A1 and T2, but not between A1 compared to T2 (likelihood ratio tests comparing the full models including all context variables with a model comprising only the autocorrelation term and the random effect; Tables 1, 3). It appeared that T4, compared to T2, was significantly less likely to occur during resting as compared to feeding and in the presence of another weaned individual (Table 4). Furthermore, T4, compared

**Table 3** Summary of results concerning contextual differences between unit types used singularly and their combinations

Comparisons	Activity		Individual within 5 m		Infant within 5 m		Vocal role	
	Resting	Foraging	Yes	No	Yes	No	Spontaneous	Reply
T4 versus T2		+		+				
T4 versus A1			+WG	+WG	+	+		
T2_T4 versus T2							+	+
T2_T4 versus T4	+				+		+	+
A1_T4 versus T4	+						+	+
A1_T4 versus A1							+	+

“+” indicates the contexts in which the respective unit types and their combinations were likely to occur

“WG” indicates contextual differences between T4 and A1 in the western gorilla group (Fig. 3). Note that while T4 was more likely in a foraging context, situations of low proximity to other individuals, A1 and T2 as well as T2\_T4 and A1\_T4 were more likely during resting and in close proximity to other individuals. Furthermore, combinations were overall more likely to be used as reply calls compared to their components used singularly. Comparison between A1 and T2, as well as between T2\_T4 and A1\_T4, was not significant (Table 1; Fig. 2)

**Table 4** Results of final model comparing the contexts of unit types T4 and T2 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate = 0.49, SE = 2.17,  $p = 0.819$ ), and two-way interactions (activity and individual within 5 m: estimate = -0.86, SE = 0.77,  $p = 0.266$ ; activity and group: estimate = 2.10, SE = 1.30,  $p = 0.105$ ; individual within 5 m and group: estimate = -1.66, SE = 1.00,  $p = 0.097$ )

	Estimate	SE	$p$ value
<i>Comparison 1: T4 (1) versus T2 (0)</i>			
Intercept	4.18	0.85	<0.001
Activity (resting)	-1.10	0.38	0.003
Individual within 5 m (yes)	-1.05	0.39	0.008
Group (western gorilla)	-1.20	0.96	0.210
Infant within 5 m (yes)	-0.57	0.37	0.126
Vocal role (spontaneous)	-0.44	0.37	0.232
Autocorrelation term	0.92	0.14	<0.001

to A1, was less likely to be emitted in the presence of an infant. The influence of the presence of other weaned individuals was different between the two study groups as indicated by a significant interaction between individuals within 5 m and group (Table 5). While in the mountain gorilla group, the presence of other weaned individuals had no obvious effects on the probabilities of T4 and A1, in western gorillas T4 was more likely when no one was present within 5 m (Fig. 3).

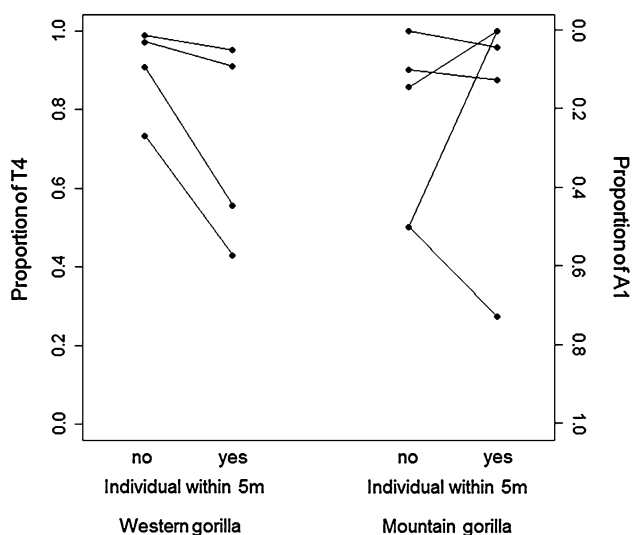
**Table 5** Results of final model comparing the contexts of unit types T4 and A1 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate = -1.74, SE = 1.66,  $p = 0.294$ ), and the nonsignificant two-way interactions (activity and individual within 5 m: estimate = 0.03, SE = 0.83,  $p = 0.974$ ; activity and group: estimate = -1.18, SE = 0.88,  $p = 0.181$ )

	Estimate	SE	$p$ value
<i>Comparison 2: T4 (1) versus A1(0)</i>			
Intercept	2.07	0.76	0.006
Activity (resting)	-0.55	0.43	0.203
Individual within 5 m (yes)	0.90	0.69	0.193 <sup>a</sup>
Group (western gorilla)	1.79	1.04	0.084 <sup>a</sup>
Individual within 5 m*Group	-2.24	0.86	0.009
Infant within 5 m (yes)	-1.37	0.66	0.037
Vocal role (spontaneous)	0.16	0.39	0.684
Autocorrelation term	0.58	0.14	<0.001

<sup>a</sup> Not meaningful in the presence of the significant interaction

Comparisons of unit types combined with others and used singularly

Overall, we found contextual differences in the probabilities of unit types used singularly compared to their combinations, but not between the combinations T2\_T4 and A1\_T4 (likelihood ratio tests, Tables 1, 3). We found that



**Fig. 3** Effect of the presence of weaned individuals within 5 m of the caller (yes/no) on the proportion of T4 as compared to A1 given by the individuals of the two study groups. *Black dots* represent the individuals of the two groups. While the presence of weaned individuals had no influence on the probabilities of T4 and A1 in mountain gorillas, in individuals of the western gorilla group T4, as opposed to A1, was more likely when no other weaned individual was in close proximity

**Table 6** Results of final model comparing the contexts of the combination T2\_T4 and unit type T2 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate = -0.17, SE = 2.24, *p* = 0.915), and nonsignificant two-way interactions (activity and individual within 5 m: estimate = -1.91, SE = 0.85, *p* = 0.16; individual within 5 m and group: estimate = -1.58, SE = 1.27, *p* = 0.214; activity and group: estimate = 1.06, SE = 1.33, *p* = 0.426)

	Estimate	SE	<i>p</i> value
<i>Comparison 4: T2_T4 (1) versus T2 (0)</i>			
Intercept	2.61	0.81	0.001
Activity (resting)	-0.24	0.38	0.533
Individual within 5 m (yes)	-0.32	0.41	0.431
Group (western gorilla)	-1.46	0.92	0.11
Infant within 5 m (yes)	0.25	0.42	0.554
Vocal role (spontaneous)	-1.41	0.40	<0.001
Autocorrelation term	0.51	0.17	0.002

T2\_T4, compared to T2, was less likely to be given spontaneously (Table 6). T2\_T4, compared to T4, was also less likely to be given spontaneously, more likely during resting and in the presence of an infant within 5 m, and furthermore, less likely to be given by the individuals of the western gorilla group (Table 7). Similarly, A1\_T4, compared to T4, was less likely to be given spontaneously, more likely during resting, and less likely to be given by the western gorilla individuals (Table 8). Moreover, A1\_T4, compared to A1, was less likely to be given

**Table 7** Results of final model comparing the contexts of the combination T2\_T4 and unit type T4 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate = 0.37, SE = 0.89, *p* = 0.683), and the nonsignificant two-way interactions (activity and individual within 5 m: estimate = 0.39, SE = 0.44, *p* = 0.382; individual within 5 m and group: estimate = -0.04, SE = 0.45, *p* = 0.924; activity and group: estimate = 0.27, SE = 0.45, *p* = 0.549)

	Estimate	SE	<i>p</i> value
<i>Comparison 5: T2_T4 (1) versus T4 (0)</i>			
Intercept	-0.41	0.35	0.242
Activity (resting)	0.56	0.29	0.053
Individual within 5 m (yes)	0.36	0.23	0.117
Group (western gorilla)	-1.13	0.54	0.037
Infant within 5 m (yes)	0.58	0.22	0.009
Vocal role (spontaneous)	-0.77	0.22	<0.001
Autocorrelation term	0.36	0.09	<0.001

**Table 8** Results of final model comparing the contexts of the combination A1\_T4 and unit type T4 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate = 1.69, SE = 1.20, *p* = 0.159), and nonsignificant two-way interactions (activity and individual within 5 m: estimate = -0.23, SE = 0.54, *p* = 0.667; individual within 5 m and group: estimate = -0.23, SE = 0.58, *p* = 0.699; activity and group: estimate = 0.43, SE = 0.56, *p* = 0.436)

	Estimate	SE	<i>p</i> value
<i>Comparison 6: A1_T4 (1) versus T4 (0)</i>			
Intercept	-0.72	0.42	0.082
Activity (resting)	0.71	0.36	0.047
Individual within 5 m (yes)	0.29	0.29	0.315
Group (western gorilla)	-1.24	0.63	0.048
Infant within 5 m (yes)	0.24	0.27	0.383
Vocal role (spontaneous)	-0.88	0.27	0.001
Autocorrelation term	0.49	0.10	<0.001

spontaneously (Table 9). Finally, T2\_A1 was less likely to be given by the western gorilla individuals compared to T2 (Table 10).

Sequential order and repetitions

Our results indicated no contextual differences in the probability of combinations with reversed order or unit types used singularly compared to combinations containing their repetition (likelihood ratio tests; Tables 1, 3).

Discussion

Previously, Hedwig et al. (2014) demonstrated that gorillas combine acoustic unit types non-randomly. As expected for



**Table 9** Results of final model comparing the contexts of the combination A1\_T4 and unit type A1 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate =  $-0.85$ , SE =  $1.86$ ,  $p = 0.649$ ), and the nonsignificant two-way interactions (activity and individual within 5 m: estimate =  $-0.11$ , SE =  $0.90$ ,  $p = 0.901$ ; individual within 5 m and group: estimate =  $-1.83$ , SE =  $0.95$ ,  $p = 0.054$ ; activity and group: estimate =  $-0.56$ , SE =  $0.92$ ,  $p = 0.547$ )

	Estimate	SE	$p$ value
<i>Comparison 7: A1_T4 (1) versus A1 (0)</i>			
Intercept	1.61	0.53	0.002
Activity (resting)	0.66	0.42	0.113
Individual within 5 m (yes)	$-0.69$	0.45	0.124
Group (western gorilla)	$-1.02$	0.64	0.109
Infant within 5 m (yes)	0.21	0.44	0.640
Vocal role (spontaneous)	$-1.10$	0.50	0.029
Autocorrelation term	0.51	0.19	0.008

**Table 10** Results of final model comparing the contexts of the combination T2\_A1 and unit type T2 used singularly after the removal of the nonsignificant three-way interaction between activity, individual within 5 m and group (estimate =  $-0.69$ , SE =  $2.57$ ,  $p = 0.786$ ), and nonsignificant two-way interactions (activity and individual within 5 m: estimate =  $-0.37$ , SE =  $1.05$ ,  $p = 0.726$ ; individual within 5 m and group: estimate =  $0.59$ , SE =  $1.60$ ,  $p = 0.711$ ; activity and group: estimate =  $-0.16$ , SE =  $1.46$ ,  $p = 0.914$ )

	Estimate	SE	$p$ value
<i>Comparison 8: T2_A1 (1) versus T2 (0)</i>			
Intercept	2.75	0.88	0.002
Activity (resting)	$-0.87$	0.49	0.077
Individual within 5 m (yes)	$-0.61$	0.49	0.208
Group (western gorilla)	$-2.86$	0.87	0.001
Infant within 5 m (yes)	0.71	0.50	0.150
Vocal role (spontaneous)	$-0.87$	0.49	0.077
Autocorrelation term	$-0.15$	0.20	0.473

syntactic systems, here we found that the combinations of unit types occurred in contexts differing from those of unit types used singularly. While the atonal grunt A1 and tonal grunt T2 used singularly appeared to be more likely during resting and in close proximity to other individuals, the grumble T4 used singularly was more likely during foraging and when individuals were alone. The combinations A1\_T4 as well as T2\_T4 occurred most likely in situations resembling the contexts of A1 and T2 used singularly, but were also more likely to be given as reply calls, whereas all three unit types used singularly were given more likely spontaneously. We found no contextual differences between A1 and T2 used singularly and combinations consisting of repetitions of them or between combinations of them given in the two different orders. This is likely due to the fact that we found no differences between A1 and T2

used singularly with regard to the contexts in which they appeared. It is possible that the distinction between atonal and tonal grunts may not be relevant to the gorillas, or it may require a more detailed analysis of acoustic or contextual variation to detect such effects.

What do these results tell us concerning the three goals we pursued in this study? Our first goal was to investigate whether gorilla close calls provide indications for lexical or phonological syntax. If lexical, combinations should be compositional in the sense that the information content of combinations can be derived from the information content of each of their unit types. Hence, we expected contextual overlap between a combination and each of its unit types used singularly. However, if phonological, the combination would occur in a different context from its unit types. The combinations A1\_T4 and T2\_T4 occurred in similar contexts to A1 or T2, but not T4. In fact, the combinations occurred in a context different from any of its unit types used singularly. T2 or A1 in combination with T4 was also used during resting and when in close proximity to other individuals, but additionally as reply calls. As such, when in combination with A1 or T2, T4 appeared to have lost its independent information content and acted as a modifier of T2 or A1. Therefore, combining unit types seems to enable gorillas to increase the number of messages that can be conveyed by only a small repertoire of five unit types, which is the key benefit of syntax (Nowak et al. 2000; Hauser et al. 2002; Fitch 2005). However, our results provide little evidence for lexical syntax because combinations did not appear to be compositional as their information content cannot be derived from the information content of each of their components. Hence, it is most parsimonious to conclude that the combinations investigated in this study constitute a form of phonological syntax.

The lack of complete contextual overlap between component calls used singularly and their combinations clearly differentiates our findings from previous examples for rudimentary forms of lexical syntax in animal vocalizations. (1) Our finding of contextual modification does not comply with the contextual modifications put forth as indicators of lexical syntax (e.g., Robins 1984), in which, for example, alarm call combinations are produced in alleviated alarm contexts (*Saguinus oedipus*: Cleveland and Snowdon 1982; *Cercopithecus diana*: Zuberbühler 2002; Ouattara et al. 2009). In contrast, the combinations investigated here were associated with a contextual modification for only one unit type (A1 or T2), whereas the context of the other component (T4) was not reflected in the use of the combination. (2) Our results neither provide an indication that the information content of calls was combined in a purely additive fashion in that compound calls are used in situations when the contexts of component calls are co-occurring (e.g., *Pan troglodytes*: Crockford and

Boesch 2005). Instead, in our study, combining was associated with a context (used as a reply call), which was unspecific to any of the component calls when used singularly. As such, our findings on gorilla close calls differ from the findings of a previous study, which suggests that the close call combinations in Diana monkeys function additively because combinations comprised a call acting as individual identifier and a call providing information about the particular context (Candiotti et al. 2012). Interestingly, our findings also show some deviation from previous examples for phonological syntax. Marler (1977) defined phonological syntax twofold. (1) It encompasses combinations of sounds without independent informational content, as found in the often stereotyped sequences of gibbons (*Hylobates agilis*: Mitani and Marler 1989). In gorillas, however, all unit types seem to have an independent information content as all of them are used singularly. Moreover, gorilla close call combinations are not produced in a stereotyped fashion but occur in great diversity as they are produced in a flexible but yet non-random way (Hedwig et al. 2014). (2) Marler's (1977) definition also included the combination of independently used sounds which lose their original information content. However, in the investigated gorilla close call combinations, only one appeared to have lost its original context (T4), while the context of the other unit types (A1, T2) was reflected by the contextual use of the compound call. However, the phonological syntax we suggest here to underlie the combining of independently used unit types in gorillas may in fact provide an intriguing analogy to linguistic phenomena since also in human language independently meaningful words used to construct sentences (lexical syntax) are frequently used as syllables to construct new words with new meaning (phonological syntax), with their original information content either being lost (e.g., "to prod" in "prod-igy") or maintained (e.g., "egg" in "egg-nog").

Our study investigated the contextual use of only a few combinations within the large syntactic variation in the gorilla close calls (see Hedwig et al. 2014). In order to fully comprehend the functioning of gorilla syntax, two main suggestions arise for further studies. (1) More combinations need to be investigated in order to gain a contextual description of syntactic variation as comprehensive as available for chimpanzees (Crockford and Boesch 2005). In chimpanzees, evidence for lexical syntax was found for some but not all call combinations (Crockford and Boesch 2005), which may be not unexpected assuming an evolutionary scenario in which lexical syntax gradually became a part of communication within the human lineage (e.g., Jackendoff 1999). (2) Based on such thorough contextual analyses, playback experiments need to be designed to explore the meaning of calls and their combinations by investigating how listeners react to specific single calls and

how this changes when calls are combined. Putty nosed monkeys, for example, seem to perceive their compound calls holistically, not compositionally (e.g., Arnold and Zuberbühler 2012).

Our second goal was to investigate possible functions of the gorilla close calls. We found that the investigated combinations occurred in two very different overall contexts, suggesting different functions. T4 appeared to be particularly emitted when foraging and when alone, both situations in which the coordination of group movements is likely to be crucial, either to cohere group members to a specific feeding site or maintain spacing while feeding, as well as to maintain contact when group members are spread out. As such, following our predictions, these results suggest that the grumble T4 may act as a signal of location. The acoustic structure of T4 appears particularly appropriate for a signal of location since its longer duration compared to T2 and A1 increases its detectability (Klump and Maier 1990; Nemeth et al. 2006). However, future studies are needed to explore the possibility that the acoustic variation within and between gorilla call combinations is adaptive in order to enhance signal transmission (Hedwig et al. in prep). In contrast, T2, A1 as well as A1\_T4 and T2\_T4 seemed more likely to occur when in close proximity to other individuals as well as during resting, when affiliative social interactions usually take place. The presence of an infant being an apparent trigger for the emission of these calls indicated a benign state of the callers, suggesting a role in facilitating affiliative social interactions, similar to the close-distance calls of various other primate species (*Saimiri sciureus*: Biben et al. 1986; *Macaca fuscata*: Masataka 1989; *Macaca arctoides*: Bauers and de Waal 1991; Bauers 1993; *Papio cynocephalus ursinus*: Cheney et al. 1995; *Macaca mulatta*: Whitham et al. 2007). Overall, our findings regarding the function of these calls are based on a rather crude contextual analysis, and more a detailed analysis is necessary. Moreover, considering that silverbacks play an important role in the gorilla social system, it is possible that the same vocalization might serve slightly different functions depending on the sex of the caller. As stated above, playback experiments investigating how listeners react to calls are ultimately needed in order to fully understand the function of these calls.

One finding of our study was that the combinations A1\_T4 and T2\_T4 were especially likely to be given as reply calls. This is particularly interesting since vocal exchanges constitute a significant aspect of gorilla vocal behavior (e.g., Harcourt et al. 1993; Harcourt and Stewart 1996). Such temporarily associated calling behavior is common among primate species and has mainly been described for long- and short-distance contact calls, which may signal identity and location when out of visual contact (e.g., *Cebuella pygmaea*: Snowdon and Cleveland 1984;

*Saimiri sciureus*: Biben et al. 1986; *Macaca fuscata*: Mitani 1986; *Saguinus oedipus*: Ghazanfar et al. 2002). Acoustic modification of reply contact calls appears to be infrequent and has been suggested to confirm the reception of the signal to a specific caller and allow for a directional exchange of information, enhance locatability of the response call (*Macaca fuscata*: Sugiura 1993, 1998), or to terminate the exchange of contact calls (*Saimiri sciureus*: Biben et al. 1986). Corroborating previous studies, our results suggest that the employment of acoustically distinct reply calls may be a relevant component of gorilla vocal behavior, which may go beyond exchanging information about the group members' locations. In this study, the reply call was most likely given during resting and when being close to other group members, which supports our hypothesis that these calls serve to coordinate affiliative interactions. A previous study reported acoustically distinct variants of double grunts in mountain gorillas used either as reply or spontaneously (Seyfarth et al. 1994), which may function to coordinate competitive interactions or the initiation of group movements (Seyfarth et al. 1994; Stewart and Harcourt 1994). Moreover, western gorillas reply with numerous close calls to a long-distance contact call (i.e., hoot series; Salmi and Doran-Sheehy 2014; Hedwig unpublished data). Given the increasing number of studies revealing high degrees of social cognition in ape vocal behavior (e.g., *Pan troglodytes*: Crockford et al. 2004; Slocombe and Zuberbühler 2005; Crockford et al. 2012; Gruber and Zuberbühler 2013; Schel et al. 2013; *Pan paniscus*: Clay and Zuberbühler 2011; *Pongo spec.*: Wich et al. 2012, but see Owren et al. 2011), the employment of such acoustically different reply calls with perhaps different informational content in gorillas may reflect a particularly flexible and expandable comprehension of vocal signals in great apes (e.g., Seyfarth and Cheney 2010). We believe that the analysis of the outcome of vocal exchanges (see Harcourt and Stewart 1996) on a detailed acoustic and contextual level will be key to testing functional hypotheses generated by more general contextual studies like the one presented here.

Additionally, we aimed to test for contextual differences in the same vocalizations used by individuals of the mountain and western gorilla groups. In both groups, the presence of an infant had a significant positive effect on the emission of A1, as opposed to T4, whereas the presence of a weaned individual only had a positive effect on the use of A1 as opposed to T4 in the western gorilla group. This is particularly interesting since western gorillas seem to show larger group spread compared to mountain gorillas (e.g., Doran and McNeilage 2001; Goldsmith 2003; Bermejo 2004; Hedwig unpublished data). Therefore, the selective pressure on the evolution of signals of location, such as our results suggest for T4,

may have been particularly strong for western gorillas. As such, our findings support a previous study by Salmi et al. (2013) suggesting contextual differences in the use of close calls in mountain and western gorillas. However, Salmi et al.'s (2013) findings indicate a more context-specific use of vocalizations in western gorillas compared to mountain gorillas due to the more restricted visibility in their lowland rainforest habitat. In contrast, our results suggest that due to their larger group spread, western gorillas may use T4 as a general signal of location as it was elicited by the absence of both infants and weaned individuals, whereas in the cohesive mountain gorillas, T4 may function as a signal of location specifically triggered by the absence of an infant. However, species differences should be inferred with caution given that our comparison was based on only two groups and a small number of individuals. Further comparative studies (based on more social groups) investigating the contextual correlates of the syntactic variation within and between species need to be carried out in order to reliably assign differences in the vocal behavior to differences in the two species' social behavior (e.g., Doran and McNeilage 2001) or habitat characteristics (e.g., Nkurunungi et al. 2005).

In conclusion, we investigated contextual correlates of the syntactic variation within gorilla close calls, which previously appeared rather enigmatic due to apparently unspecific contextual use (e.g., Harcourt et al. 1993; Salmi et al. 2013). However, by using a contextually rather crude level of analysis, we were able to elucidate potential functions of different gorilla close calls, suggesting that the syntactic approach followed here may be a fruitful framework for future studies. Based on our results, we hypothesize that the combination of acoustic units into compound vocalizations by gorillas constitutes a form of phonological syntax, which may enable them to increase the number of messages that can be transmitted by an otherwise small repertoire of acoustic units. However, future studies are needed to test this hypothesis by including more combinations into the analysis and particularly using playback experiments investigating how gorillas perceive the unit types used singularly and when combined. Together with the few previous studies investigating syntactic structures in ape vocal communication (Crockford and Boesch 2005; Clay and Zuberbühler 2011; Lameira et al. 2013), our results provide a baseline for future studies aiming to achieve a more comprehensive understanding of great ape vocal communication and their syntactic capabilities, which is crucial to our understanding of the evolutionary origins of human language.

**Acknowledgments** 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 humor, and courage. This research was financially supported by the Leakey Foundation and the Max Planck Society.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Altmann J (1974) Observational study of behaviour: sampling methods. *Behaviour* 49:227–267
- Arnold K, Zuberbühler K (2012) Call combinations in monkeys: compositional or idiomatic expressions? *Brain Lang* 120:303–309
- Baayen RH (2008) *Analysing Linguistic Data*. Cambridge University Press, Cambridge
- Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: keep it maximal. *J Mem Lang* 68:255–278
- Bates D, Maechler M, Bolker B, Walker S (2013) *Lme4: linear mixed-effects models using Eigen and S4*. R package version 1.0-4
- Bauers KA (1993) A functional analysis of staccato grunt vocalizations in the stump-tailed macaque (*Macaca arctoides*). *Ethology* 94:147–161
- Bauers KA, de Waal FB (1991) “Coo” vocalizations in stump-tailed macaques: a controlled functional analysis. *Behaviour* 119:143–160
- Bermejo M (2004) Home-range use and intergroup encounters in western gorillas (*Gorilla g. Gorilla*) at Lossi Forest, North Congo. *Am J Primatol* 64:223–232
- Biben M, Symmes D, Masataka N (1986) Temporal and structural analysis of affiliative vocal exchanges in squirrel monkeys (*Saimiri sciureus*). *Behaviour* 98:259–273
- Boinski S (1993) Vocal coordination of troop movement among white-faced capuchin monkeys *Cebus capuchinus*. *Am J Primatol* 30:85–100
- Boinski S, Campbell AF (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 Psych* 126:45–56
- Candiotti A, Zuberbühler K, Lemasson A (2012) Context-related call combinations in female Diana monkeys. *Anim Cogn* 15:327–339
- Cheney DL, Seyfarth RM, Silk JB (1995) The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Anim Behav* 50:249–257
- Chomsky N (1965) *Aspects of the theory of syntax*. MIT press, Cambridge
- Clay Z, Zuberbühler K (2011) Bonobos extract meaning from call sequences. *PLoS ONE* 6:e18786
- Cleveland J, Snowdon CT (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 RM, Mundry R, Zuberbühler K (2012) Wild chimpanzees inform ignorant group members of danger. *Curr Biol* 22:142–146
- Dobson AJ (2002) *An introduction to generalized linear models*. CRC Press, Baton Rouge
- Doran DM, McNeilage A (2001) Subspecific variation in gorilla behaviour: the influence of ecological and social factors. In: Robbins MM, Sicotte P, Stewart KJ (eds) *Mountain gorillas: three decades of research at Karisoke*. Cambridge University Press, New York, pp 123–149
- Elowson AM, Snowdon CT (1994) Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Anim Behav* 47:1267–1277
- Field AP (2005) *Discovering statistics using SPSS*. SAGE, London
- Fitch WT (2005) The evolution of language: a comparative review. *Biol Philos* 20:193–203
- Forstmeier W, Schielzeth H (2011) Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner’s curse. *Beh Ecol Sociobiol* 65:47–55
- Fossey D (1972) Vocalizations of the mountain gorilla (*Gorilla beringei*). *Anim Behav* 20:36–53
- Fox J, Weisberg S (2011) *An R companion to applied regression*, 2nd edn. Thousand Oaks, CA
- Ghazanfar AA, Smith-Rohrberg D, Pollen AA, Hauser MD (2002) Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins. *Anim Behav* 64:427–438
- Goldsmith ML (2003) Comparative behavioural ecology of a lowland and highland gorilla population: where do Bwindi gorillas fit? In: Hohmann G, Robbins MM, Boesch C (eds) *Feeding ecology in apes and other primates: ecological, physical and behavioural aspects*. Cambridge University Press, New York, pp 358–384
- Gruber T, Zuberbühler K (2013) Vocal recruitment for joint travel in Wilde Chimpanzees. *PLoS ONE* 8:e76073
- Harcourt AH, Stewart KJ (1996) Function and meaning of wild gorilla ‘close’ calls 2. Correlations with rank and relatedness. *Behaviour* 133:827–845
- Harcourt AH, Stewart KJ (2001) Vocal relationships of wild mountain gorillas. In: Robbins MM, Sicotte P, Stewart KJ (eds) *Mountain gorillas: three decades of research at Karisoke*. Cambridge University Press, New York, pp 241–262
- Harcourt AH, Stewart KJ, Harcourt DE (1986) Vocalizations and social relationships of wild gorillas: a preliminary analysis. In: Taub DM, King FA (eds) *Current perspectives in primate social dynamics*. Van Nostrand Reinhold Co., New York, pp 346–356
- Harcourt AH, Stewart KJ, Hauser M (1993) Functions of wild gorilla ‘close call’s. I. Repertoire, context, and interspecific comparison. *Behaviour* 124:89–122
- Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298:1569–1579
- Hedwig D, Hammerschmidt K, Mundry R, Robbins MM, Boesch C (2014) Acoustic structure and variation in mountain and western gorilla close calls: a syntactic approach. *Behaviour* 151:1091–1120
- Hurford JR (2011) *The origins of grammar: language in the light of evolution II*. Oxford Studies in the evolution of language. Oxford University Press, Oxford
- Jackendoff R (1999) Possible stages in the evolution of the language capacity. *Trends Cogn Sci* 3:272–279
- Kanwal JS, Matsumura S, Ohlemiller K, Suga N (1994) Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. *J Acoust Soc Am* 96:1229–1254

- Kershenbaum A, Ilany A, Blaustein L, Geffen E (2012) Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proc R Soc A* 279:2974–2981
- Klump GM, Maier EH (1990) Temporal summation in the European starling (*Sturnus vulgaris*). *J Comp Psych* 104:94–100
- Lameira AR, de Vries H, Hardus ME, Hall CP, Mitra-Setia T, Spruijt BM, Wich SA (2013) Predator guild does not influence orangutan alarm call rates and combinations. *Behav Ecol Sociobiol* 67:519–528
- 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
- Marler P (1977) The structure of animal communication sounds. In: Bullock TH (ed) *Recognition of complex acoustic signals: report of the Dahlem workshop*. Abakon Verlagsgesellschaft, Berlin
- Marler P, Mitani J (1988) Vocal communication in primates and birds: parallels and contrasts. In: Todt D, Goedeck P, Symmes D (eds) *Primate vocal communication*. Springer, Berlin, pp 1–14
- Masataka N (1989) Motivational referents of contact calls in Japanese monkeys. *Ethology* 80:265–273
- Masi S, Cipolletta C, Robbins MM (2009) Western lowland gorillas (*Gorilla gorilla*) change their activity patterns in response to frugivory. *Am J Primatol* 71:91–100
- Mitani M (1986) Voiceprint identification and its application to sociological studies of wild Japanese monkeys (*Macaca fuscata ayaku*). *Primates* 27:397–412
- Mitani JC, Marler P (1989) A phonological analysis of male gibbon singing behavior. *Behaviour* 109:20–45
- Nemeth E, Dabelsteen T, Pedersen SB, Winkler H (2006) Rainforests as concert halls for birds: are reverberations improving sound transmission of long song elements? *J Acoust Soc Am* 119:620–626
- Nkurunungi JB, Ganas J, Robbins MM, Stanford CB (2005) A comparison of two mountain gorilla habitats in Bwindi Impenetrable National Park, Uganda. *Afr J Ecol* 42:289–297
- Nowak MA, Plotkin JB, Jansen VA (2000) The evolution of syntactic communication. *Nature* 404:495–498
- Ouattara K, Lemasson A, Zuberbühler K (2009) Campbell's monkeys use affixation to alter call meaning. *PLoS ONE* 4:e7808
- Owren MJ, Amoss RT, Rendall D (2011) Two organizing principles of vocal production: implications for Nonhuman and Human Primates. *Am J Primatol* 73:530–544
- Portner PH (2005) *What is meaning? Fundamentals of formal semantics*. Blackwell, Malden
- R Core Development Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria
- Robbins MM (2008) Feeding competition and agonistic relationships among Bwindi Gorilla beringei. *Int J Primatol* 29:999–1018
- Robbins MM (2010) Gorillas: diversity in ecology and behaviour. In: Campbell CJ, Fuentes A, MacKinnon KC, Bearder S, Stumpf R (eds) *Primates in perspective*. Oxford University Press, Oxford, pp 326–339
- Robinson JG (1984) Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus olivaceus*. *Behaviour* 90:46–79
- Salmi R, Doran-Sheehy DM (2014) The function of loud calls (Hoot Series) in wild western gorillas (*Gorilla gorilla*). *Am J Phys Anthropol*. doi:10.1002/ajpa.22575
- Salmi R, Hammerschmidt K, Doran-Sheehy DM (2013) Western gorilla vocal repertoire and contextual use of vocalizations. *Ethology* 119:831–847
- Schaller GB (1963) *The mountain gorilla*. University of Chicago Press, Chicago
- Schassburger RM (1993) *Vocal communication in the timber wolf, Canis lupus, Linnaeus: structure, motivation and ontogeny*. Paul Parey Scientific Publishers, Berlin
- Schel AM, Machanda Z, Townsend SW, Zuberbühler K, Slocombe KE (2013) Chimpanzee food calls are directed at specific individuals. *Anim Behav* 86:955–965
- Schielzeth H, Forstmeier W (2009) Conclusions beyond support: overconfident estimates in mixed models. *Behav Ecol* 20:416–420
- Seyfarth RM, Cheney DL (2010) Production, usage, and comprehension in animal vocalizations. *Brain Lang* 115:92–100
- Seyfarth RM, Cheney DL, Harcourt AH, Stewart KJ (1994) The acoustic features of gorilla double grunts and their relation to behaviour. *Am J Primatol* 33:31–50
- Silk JB, Kaldor E, Boyd R (2000) Cheap talk when interests conflict. *Anim Behav* 59:423–432
- Slocombe KE, Zuberbühler K (2005) Functionally Referential Communication in a Chimpanzee. *Curr Biol* 15:1779–1784
- Snowdon CT (1997) Affiliative processes and vocal development. *Ann N Y Acad Sci* 807:340–351
- Snowdon CT, Cleveland J (1984) "Conversations" among Pygmy Marmosets. *Am J Primatol* 7:15–20
- Stephan C, Zuberbühler K (2008) Predation increases acoustic complexity in primate alarm calls. *Biol Lett* 4:641–644
- Stewart KJ, Harcourt AH (1994) Gorillas' vocalizations during rest periods: signals of impending departure? *Behaviour* 130:29–40
- Sugiura H (1993) Temporal and acoustic correlates in vocal exchange of coo calls in Japanese macaques. *Behaviour* 124:207–225
- Sugiura H (1998) Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim Behav* 55:673–687
- 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
- Whitham JC, Gerald MS, Maestripieri D (2007) Intended receivers and functional significance of grunt and girney vocalizations in free-ranging female rhesus macaques. *Ethology* 113:862–874
- Wich SA, Krutzen M, Lameira AR, Nater A, Arora N, Bastian ML, Meulman E, Morrogh-Bernard HC, Atmoko SSU, Pamungkas J, Perwitasari-Farajallah D, Hardus ME, van Noordwijk M, van Schaik CP (2012) Call cultures in orang-utans? *PLoS ONE* 7:e36180
- Zuberbühler K (2002) A syntactic rule in forest monkey communication. *Anim Beh* 63:293–299