RESEARCH PAPER

Acoustic Properties of Long Calls Given by Flanged Male Orang-Utans (*Pongo pygmaeus wurmbii*) Reflect Both Individual Identity and Context

Brigitte Spillmann*, Lynda P. Dunkel*, Maria A. van Noordwijk*, Rahmalia N. A. Amda†, Adriano R. Lameira‡, Serge A. Wich§ & Carel P. van Schaik*

* Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland

† Fakultas Biology, Universitas Nasional, Jakarta, Indonesia

‡ Behavioural Biology Group, Utrecht University, Utrecht, The Netherlands

§ Great Ape Trust of Iowa, Des Moines, IA, USA

Correspondence

Brigitte Spillmann, Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland.

E-mail: spillbri@gmx.ch

Abstract

Long calls by flanged male Bornean orang-utans (*Pongo pygmaeus wurmbii*) serve as a long-distance communication signal in this semi-solitary species and allow individuals to adjust their ranging behavior. Long calls can be heard up to circa 1 km in dense rainforest. Only flanged males emit them, in various contexts: spontaneously (where no disturbances from the environment are perceived by human observers), when highly aroused by another male's long call or a falling tree nearby, or right after having pushed over a dead tree themselves. In this study, acoustic analyses of orang-utan long calls at Tuanan in Central Kalimantan not only confirm the discrimination of individual males by their long calls but also demonstrate the discrimination of context based on the long calls' acoustic structure, which is further supported by the females' ranging responses according to long call contexts. Females with dependent offspring move away from spontaneous long call sources but appear to ignore long calls elicited by disturbance. Hence, Bornean orang-utan females perceive measurable differences in acoustic characteristics of long calls given in different contexts. These findings concur with vocal discrimination of contexts in other non-human primates.

Introduction

Long distance vocalizations are widespread in many organisms living in habitats with limited visibility, including non-human primates. Although the functions of these calls may vary (Zuberbühler et al. 1997; Wich & Nunn 2002; da Cunha & Byrne 2006), they would generally be expected to reveal the identity of the caller and perhaps details of the caller's internal state, as affected by the call's context. Indeed, in many non-human primate species, individual identity is encoded in both short- and long-distance vocalizations (Waser 1977; Hammerschmidt & Todt 1995; Mitani et al. 1996; Teixidor & Byrne 1999; Wich et al. 2003; Rendall et al. 2004). A particular vocalization of an individual may also differ depending on the social or the environmental context in which it is given in a variety of primate species (Zuberbühler et al. 1997; Norcross et al. 1999; Crockford & Boesch 2003; Wich et al. 2003; Notman & Rendall 2005), thus suggesting an effect of the caller's internal state (Hauser 1996). For instance, Wich et al. (2003) have shown that the reactions of female Thomas' langurs (*Presbytis thomasi*) to male loud calls depended on context, suggesting that the audience makes use of the variation in acoustic properties of the loud calls.

Long-term studies at a range of sites have revealed that orang-utans have a fairly rich repertoire of calls (vocalizations and non-vocal sounds) (Hardus et al. 2009). However, only sexually mature, flanged males emit long-distance vocalizations, the so-called long calls. Long calls are a series of loud booming vocalizations called pulses, generally preceded by a softer vocalization called grumbles, and often ending with multiple bubbles. The pulses can be heard over 1 km, depending on terrain. Individual males emit long calls approximately four times a day during the active day, and at an unknown (but lower) rate during the night (Delgado et al. 2009).

Individual distinctiveness in orang-utan long calls was demonstrated in previous studies by acoustic analyses. Behavioral observations, including playback experiments also indicate individual recognition of the callers (Mitani 1985b; Delgado 2003; Mitra Setia & van Schaik 2007; Lameira & Wich 2008). However, males emit long calls in multiple contexts (Galdikas 1983). Firstly, many long calls are emitted spontaneously without any stimuli from the environment detectable to human observers. Because no unusual behaviors are observed before the call is given, the caller does not seem to be especially aroused. Secondly, males may spontaneously emit long calls after having given displays indicative of high arousal, such as snag crashing (the pushing over of dead trees). Finally, long calls may be given in response to a social or environmental disturbance (i.e. in response to another male's long call, after a tree fall, which often indicates snag crashing by an aroused rival or observers). The agitated behavior of the males in this context also indicated high arousal. To date, the possible effects of these context differences on acoustic structure have not been examined.

In this study, we analyzed acoustic properties of long calls to examine whether individual identity and context are both encoded in a long call. We also examined whether the audience used the context information by examining females' ranging responses in relation to long call context. In the discussion, we address the consequences of context effects on the acoustic structure of long calls for hypotheses about long call function.

Methods

Study Site and Behavior

The research project was conducted at Tuanan, Central Kalimantan, Borneo, Indonesia in 2005 and 2007. Tuanan (2°09′S, 114°26′E) is located near the Kapuas River in the Mawas Reserve. The research area encompasses a 750 ha grid-based trail system in a heavily logged peat swamp forest. The average orang-utan density is at least 4.25 ind/km² (van Schaik et al. 2005). The behavior of males and females was recorded during focal follows, using a standard recording protocol (basically following van Schaik (1999); see also http://www.aim.uzh.ch/orangutannetwork.html). Recordings were made during focal follows of the males.

Long Call Recordings and Contexts

Long calls of two flanged males (Niko and Kentung) were recorded using a Sennheiser ME 67 shotgun microphone (Sennheiser electronic GmbH & Co.KG, Wedemark-Wennebostel, Germany) and a Roland Edirol R-09 digital recording device (Roland Corporation, Hamamatsu, Japan) (with sampling rate set to 44100 Hz and sample size to 16-bit signed PCM). Long calls of a third flanged male (Fugit) were recorded with a Marantz Analogue Recorder PMD222 (Marantz Corporation, Kenagawa, Japan) in combination with a Sennheiser Microphone ME 64 or a Sony Digital Recorder TCD-D100 (Sony Corporation, Tokyo, Japan) in combination with a Sony Microphone ECM-M907 (Sony Corporation, Tokyo, Japan) by A. L. in 2005 (analogue recorded long calls were digitized at 44.1 kHz). The behavioral context was documented when a long call occurred and divided into (i) spontaneously emitted long calls without any disturbances observable by human observers; (ii) long calls accompanied by the male snag crashing; and (iii) long calls given in response to a disturbance, such as a long call of another male or a tree falling nearby. The latter two contexts (ii and iii) can be described as high-arousal long calls (Table 1).

Table 1: Overview of the number of recorded long calls in different contexts

Focal male	Spontaneously emitted long calls (i)	Long calls in arousal (ii and iii)	Long calls with previous snag crash (ii)	Long calls emitted by disturbance (iii)	Total
Male 1	10	5	2	3	15
Male 2	10	12	1	11	22
Male 3	2	10	5	5	12
Total	22	27	8	19	49

Acoustic Analyses

Spectrogram analyses were conducted in Raven 1.1 (Update 10.1 Copyright 2003, Cornell Lab of Ornithology Bioacoustic Research Program). Long call recordings were transformed into spectrograms (window type = Hann, window size = 1024 samples, 3dB bandwidth = 61.9 Hz, time grid spacing = 102 samples, frame overlap = 90%, frequency grid spacing = 43.1 Hz). The analyses of the long calls comprise frequency and temporal measurements, as well as measurements of descriptive variables for the long call as a whole, i.e. call rate (number of pulses/s), long call duration, number of pulses, and number of pulses per pulse type (see Table 2). Frequency and temporal measurement were taken from the fundamental frequency of the spectrogram view in Raven 1.1 after converting waveform into a spectrogram view by fast Fourier transformation (FFT).

Long calls consist of a series of loud booming vocalizations called pulses. Long calls often start with preceding 'grumbles' (bubbling-like elements) that are low in loudness, mostly followed by high amplitude pulses as a climax section and then turn into a slowed down part that can be terminated by 'bubbles.' Because pulse types vary in their acoustic properties, we rely on the classification by Ross & Geissmann (2007), who recognized five different pulse types varying in frequency modulation: roar (high amplitude pulse type), huitus (high amplitude pulse type, with steeply ascending and descending part that are not connected), intermediary (low amplitude pulse type, frequency modulation starts with a rising part followed by a falling part that changes again into a rising and ends with a falling part), sigh (low amplitude pulse type, starts with a short rising part and changes in a long falling part, and bubbles (low amplitude pulse type, looks like a cracked sigh). The pulse type 'roar' should not be confused with the vocalization called roar (Hardus et al. 2009). Pulses were generally classified following the rules of Ross & Geissmann (2007). However, because pulse type 'roar' showed a variety of frequency modulations we reclassified this original 'roar' into three subgroups: 'roar,' 'volcano' and 'low roar' (see all relevant pulse types in Fig. 1). Pulse type 'volcano' and 'roar' differ at the top of the fundamental frequency, where 'volcano' shows a sharp tip and a higher frequency (mean: 767 Hz) and pulse type roar is more rounded and lower in frequency (mean: 586 Hz) (see Fig. S1 and Table S1). 'Low roars' show circa half of the fundamental frequency at the highest point than the 'roars' and are therefore classified as a pulse type on its own.

Previous analyses have used all high-amplitude pulses produced during the first 30 s of the long call without distinguishing between the types of pulses. However, if one simply measures variables of the first n pulses, the differences in pulse type composition can easily overwhelm individual differences among males. Nonetheless, each individual male can produce each of these pulse types (Fig. 1: long call containing all relevant pulse types). To reach a comparison that is not influenced by different pulse type composition we therefore decided to limit our spectrogram analyses to the acoustic properties of only one pulse type ('roar,' narrowly defined), because this was common in long calls of all males and represent a loud, high-amplitude pulse type that
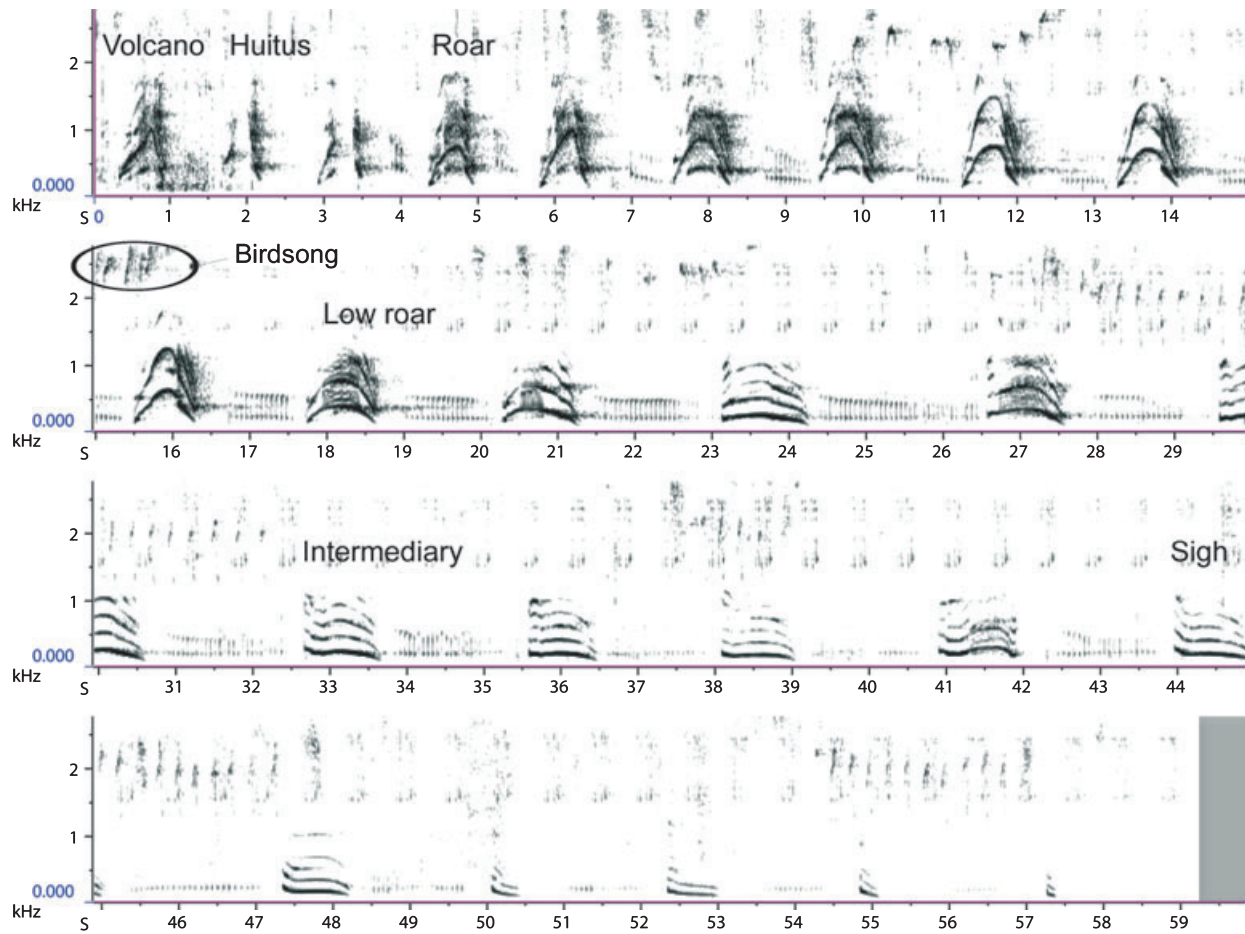
Table 2: Overview of variables measured with Raven from pulse type 'roar'. Descriptive measurements refer to the long call as a whole.

Raven 1.1 measurements		
F_0 start	[Hz]	Fundamental frequency start
F_0 high	[Hz]	Highest point of fundamental Frequency
F_0 end	[Hz]	Fundamental frequency end
F_0 rising duration	[s]	Duration of F_0 rising
F_0 rising range	[Hz]	Bandwidth of F_0 rising
F_0 falling duration	[s]	Duration of F_0 falling
F_0 falling range	[Hz]	Bandwidth of F_0 falling
Dur. pulse	[s]	Duration of pulse
Dur. interval	[s]	Duration of interval (time between two pulses)
Derived characters		Dataset 1 & 2
Rising slope	[Hz/s]	Slope of F_0 rising (F_0 rising range/F_0 rising duration)
Falling slope	[Hz/s]	Slope of F_0 falling (F_0 falling range/F_0 falling duration)
Descriptive variables		Dataset 1 & 2
Duration of long call	[s]	Duration of long call including bubbles
Pulses	#	No. pulses
Roars	#	No. pulses
Double roars	#	No. roars
Volcanoes	#	No. volcanoes
Huitus	#	No. huitus
Low roars	#	No. low roars
Intermediaries	#	No. intermediaries
Sighs	#	No. sighs
# Bubbles	#	No. bubbles
Call rate	#/[s]	No. pulses/LC duration (grumbles and bubbles were excluded)
Call rate first 30 s	#/[s]	No. pulses/first 30 s. (grumbles were excluded whether occurred)

Fig. 1: Representative long call of a flanged male (Niko). All different pulse types used in our analyses are labeled. At higher frequencies (1500–2800 Hz) a birdsong is visible.

carry far. Although this means that the acoustic properties of the other pulse types were not included in the analysis, information on their number was still included. 'Roars' mostly occur at the beginning of the long call. Whenever possible the first pulse or pulses of this particular pulse type was measured.

Statistical Analyses

Principal components and discriminant analyses were carried out using SPSS 16.0 (SPSS Inc., Chicago, IL, USA). First of all, a principal component analysis was performed for pulse type 'roar' separately to reduce the number of variables and to eliminate potential correlations among them. Principal components with eigenvalue ≥1 (Hirsig 2002) were selected and then rotated by varimax factor rotation. The principal components analyses reduced the original 11 variables to three rotated factors for spectral and temporal measurements of pulse type 'roar.'

The factor rotation values of the previous principal component analysis of pulse type 'roar' were then combined with the descriptive variables of the long calls, such as duration of long call or call rate. This data set was again subjected to principal components analysis to reduce the number of variables. The first five principal components (eigenvalue ≥1) of the analysis were rotated with the varimax factor rotation.

Discriminant function analysis was conducted on these five rotated factors (eigenvalue ≥1) to examine the possible identification of the males by their long calls and the different contexts in which long calls were given. Discriminant function analyses were conducted entering all independent factors simultaneously. Results from the analyses were cross-validated using leave-one-out classification, and statistical significance was assessed using Cohen's kappa (Titus et al. 1984).

In a first discriminant function analysis of context, long call contexts were classified into two groups.

Group 1 contained long calls given spontaneously. Spontaneous long calls were long calls without any disturbances from the environment detectable to the human observer and without previous snag crashes by the focal male. The long calling male did not seem to be aroused. Group 2 contained long calls given after disturbances from the environment or after a display (snag crash).

In a second discriminant function analysis, long call contexts were classified into three groups. The first group contains spontaneous long calls as above. The second group contains long calls elicited by disturbances (i.e. in response to another long call or a nearby tree fall) and the third group contains long calls preceded by a snag crash by the calling male where no disturbances from the environment were observed.

Females' ranging responses

Whenever we had simultaneous follows of a calling male and a female (that heard the long call but was not in association with the male), we analyzed the ranging responses of the females to a long call at 30, 60, 90, and 120 min after the long call, differentiated by long call context. The ranging responses were calculated directly from the map by subtracting the distance between the long call position and the female position after all time steps from the distance between the long calling male and the female during the long call. The females' delayed responses in terms of changes in ranging better differentiate their assessment of a long call than their immediate responses upon hearing the call (Delgado 2003; Mitra Setia & van Schaik 2007). The number or ranging responses decreases over time because of the occurrence of new long calls heard by a female within 2 hours. Ranging responses to the former long call were then excluded from that point on to maintain independence. Sexually active females were excluded from the analyses due to the combination of their having different ranging responses to long calls compared with females with dependent offspring (L. Dunkel unpubl. data) and small sample size.

Results

Five rotated factors (eigenvalue ≥ 1) explain 72.8% of the long call variability. These five rotated factors were introduced into three different discriminant function analyses. Individual discrimination on these five rotated factors was possible for 85.7% of

Table 3: Results of discriminant function analyses

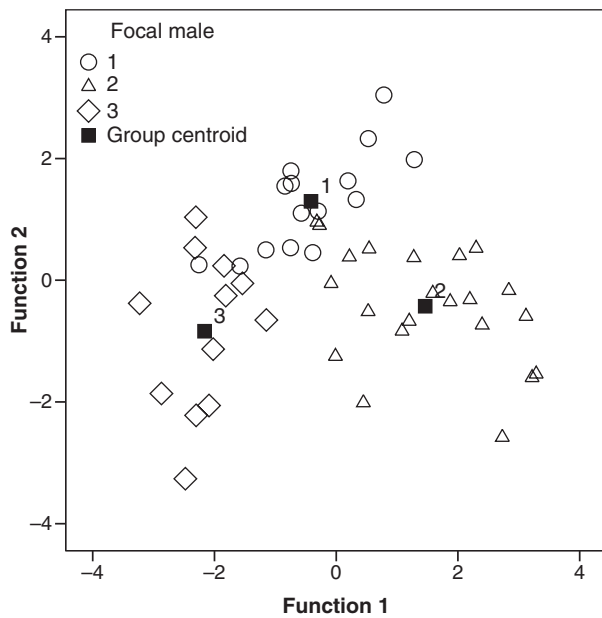| Discriminant function analysis	No. of individuals	No. of discriminant functions (% of explained variance)	Correct allocation in %	No. of misclassified long calls	Chance-corrected classification Cohen's kappa	Correct allocation with 'leave-one-out' validation in %	Chance-corrected classification Cohen's kappa	By chance in %
1. Individual dicrimination	3	2 (73.7, 26.3)	85.7	7 out of 49	$P_0 = 0.857$ $P_C = 0.338$ kappa: 0.784 $Z = 7.686$ $p < 0.001$	75.5	$P_0 = 0.755$ $P_C = 0.338$ kappa: 0.63 $Z = 6.176$ $p < 0.001$	33.33
2. Context discrimination (no arousal vs. arousal)	3	1 (100)	83.7	8 out of 49	$P_0 = 0.837$ $P_C = 0.492$ kappa: 0.679 $Z = 6.528$ $p < 0.001$	79.6	$P_0 = 0.796$ $P_C = 0.492$ kappa: 0.598 $Z = 4.241$ $p < 0.001$	50
3. Context discrimination (spontaneous, snag crash, disturbance)	3	2 (80.1, 19.9)	81.6	9 out of 49	$P_0 = 0.796$ $P_C = 0.364$ kappa: 0.679 $Z = 6.287$ $p < 0.001$	67.3	$P_0 = 0.673$ $P_C = 0.366$ kappa: 0.484 $Z = 4.44$ $p < 0.001$	33.33

Fig. 2: Scatter plot of individual scores for discrimination function 1 against discrimination function 2 from dataset 2.
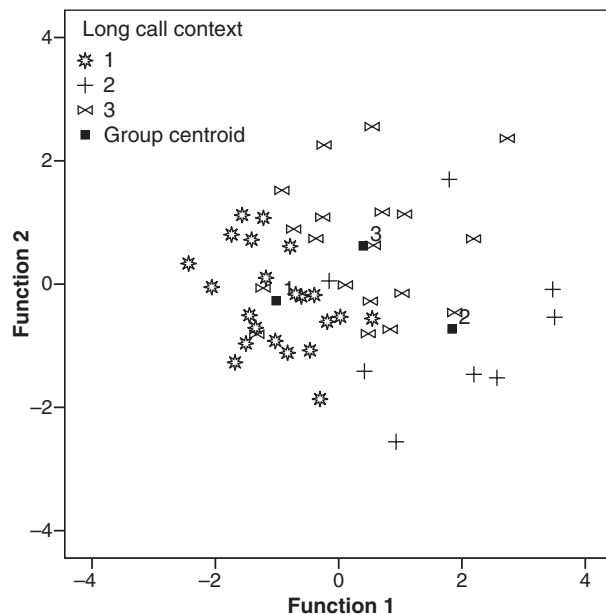

Fig. 3: Scatter plot of context scores for discriminant function 1 against discriminant function 2 from dataset 2.

the long calls (with leave-one-out method 75.5%). Discrimination of spontaneously emitted long calls and long calls emitted in an aroused affective state gave the correct allocation for 83.7% of the long calls (with leave-one-out method 79.6%). As shown in Table 3 and Figs 2 and 3, correct

allocation of three contexts was 81.6% (with leave-one-out method 67.3%). Chance-corrected classification after Cohen is significant in all discriminant function analyses.

Parameters Responsible for Discrimination

Correlations between the five rotated factors with the first and the second discriminant functions >0.45 showed that individual discrimination is characterized by a combination of frequency, temporal, and descriptive variables (frequency measurements: F_0 high, F_0 rising range, F_0 falling range, F_0 falling slope; temporal measurements: F_0 rising duration and duration of pulses; descriptive measurements: call rate, long call duration, number of pulses, number of pulse types 'sigh,' 'volcano,' and 'bubbles'). Context discrimination is explained by temporal and descriptive variables (temporal measurements: F_0 rising duration and duration of pulses; descriptive measurements: call rate, long call duration, number of pulses, number of pulse type 'huitus,' and 'bubbles'). Long calls given during high arousal differ from spontaneously given long calls in that pulse duration is shorter, the F_0 rises faster to its peak, call rate is faster, and the number of pulses is larger (Fig. 4). The best single predictor variable to differentiate the three long call contexts is call rate. Spontaneously given long calls are the slowest, long calls given in response to disturbance are intermediate and long calls following a previous snag crash are the fastest. Univariate and descriptive analyses of different long call measurements are listed in the supporting information (Figs S2–S4, Tables S2 and S3).

Audience Responses

The ranging responses of females are significantly different according to the context after 60, 90 and 120 min of the long call. Females seem to differentiate between the different contexts, regardless of whether we recognized two (Fig. 5a) or three (Fig. 5b) contexts. The differences in the females' ranging responses between no arousal (spontaneous) and high arousal (disturbance and snag crash) long calls were assessed with a Wilcoxon Test. Differences were significant 60 min after the long call ($Z = 2.275$, p = 0.0222, $n_{spont.} = 30$, $n_{arousal} = 16$), 90 min after the long call ($Z = 2.772$, p = 0.0053, $n_{spont.} = 25$, $n_{arousal} = 11$), and 120 min after the long call ($Z = 2.471$, p = 0.0126, $n_{spont.} = 20$, $n_{arousal} = 10$): females moved away from spon-
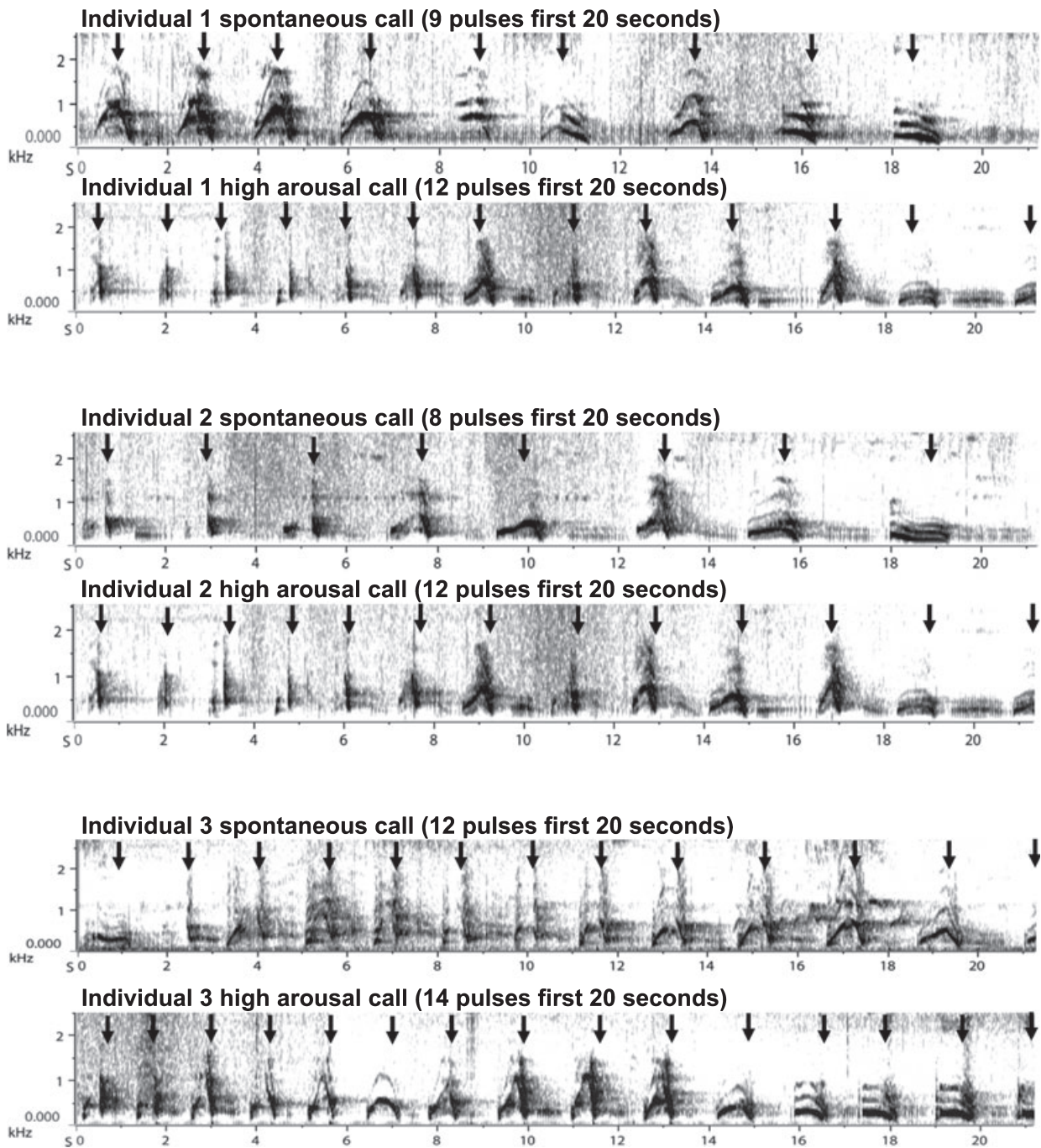
Individual 1 spontaneous call (9 pulses first 20 seconds)

Individual 1 high arousal call (12 pulses first 20 seconds)

Individual 2 spontaneous call (8 pulses first 20 seconds)

Individual 2 high arousal call (12 pulses first 20 seconds)

Individual 3 spontaneous call (12 pulses first 20 seconds)

Individual 3 high arousal call (14 pulses first 20 seconds)

Fig. 4: First 20 s of long calls given spontaneously and in high arousal from all three individuals.

taneously given long calls but more or less ignore long calls given in high arousal.

Females' ranging response also differed between three contexts (spontaneous, snag crash, and disturbance calls) as measured with a Kruskal–Wallis test, after 60 min ($\chi^2 = 7.0108$, p = 0.03, $n_{spont.} = 30$, $n_{snagcrash} = 9$, $n_{disturb.} = 7$), 90 min ($\chi^2 = 9.2228$, p = 0.0099, $n_{spont.} = 25$, $n_{snagcrash} = 7$, $n_{disturb.} = 4$), and 120 min ($\chi^2 = 6.8190$, p = 0.0331, $n_{spont} = 20$, $n_{snagcrash} = 7$, $n_{disturb.} = 3$). A *post hoc* Tukey HSD test showed significance at the 0.05 level between ranging responses after spontaneous long calls and disturbance calls (after 60 min: p = 0.039, after 90 min: p = 0.006, after 120 min: p = 0.036),
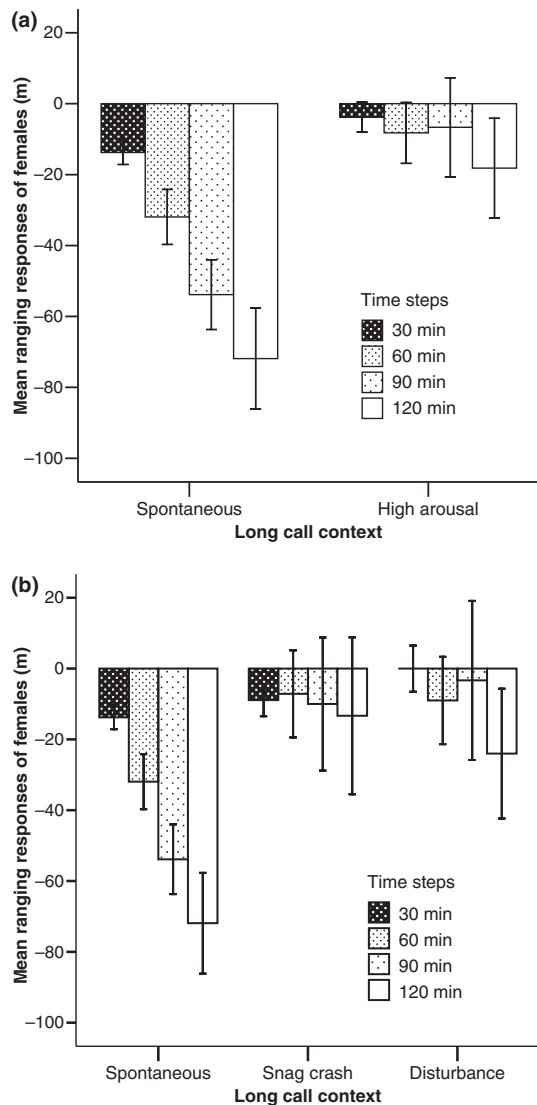
Fig. 5: (a) Females' mean ranging responses 30, 60, 90 and 120 min after a long call. Negative values indicate females moving away from the source of the call. Numbers of females ranging responses are given beyond the bars (±1 SE). (b) Females' mean ranging responses 30, 60, 90 and 120 after a long call according to the three long call contexts (±1 SE).

whereas ranging responses given after long calls with preceding snag crash were not significantly different from the other two contexts.

Discussion

Methodological Aspects

Previous studies had already performed acoustic analyses of long calls, and shown individual distinctiveness as well as geographic variation (Delgado 2007; Ross & Geissmann 2007). In this study, both individual distinctiveness and context variation were analyzed. In this analysis, we used the five types of pulses described by Ross & Geissmann (2007), but reclassified their pulse type 'roar' into two different pulse types, namely 'roar' (narrowly defined) and 'volcano' (see Fig. S1), because they were significantly different in nine out of 11 spectral and temporal measurements, especially frequency modulation (Table S1). In the discriminant function analyses presented here, we only included temporal and spectral measurements of pulse type 'roar' even though all males are able to produce both 'roars' and 'volcanoes.' However, including both types in the discriminant analysis did not change the results (not shown). Thus, it was justified to subdivide the pulse type 'roar.'

Individual and Contextual Differentiation

Individual recognition is important in long-distance communication whenever individuals are separated beyond visual contact. In Sumatra, previous studies at two different sites showed that females irrespective of reproductive state mostly approach the long calls of the dominant male in the area (Delgado 2003; Mitra Setia & van Schaik 2007). These findings suggest that there is individual recognition beyond visual contact. Analyses of call structure confirmed the feasibility of acoustic long call discrimination in this population (Delgado 2003).

Among Sumatran orang-utans, spontaneous long calls indicate the subsequent travel direction of the calling males for many hours (C. van Schaik, unpublished), whereas long calls caused by disturbances from the environment do not. The latter calls may be the result of changes in the affective state of a flanged male (Hauser 1996) triggered by a vocalization of a conspecific, a falling tree nearby, disturbance by observers, or a preceding display, such as snag crashing (Galdikas 1985a; Galdikas & Insley 1988). In this project, we were able to show for Bornean orang-utans that differences in acoustic properties of long calls are associated with context differences.

Long calls emitted in an aroused affective state are slightly faster, have pulses of shorter duration, and contain more pulses and bubbles than spontaneously emitted long calls (audiofiles of long call types can be found in the supporting information). Although arousal effects on calls of the same type can be expressed in both frequency and time domains (Schehka et al. 2007), the aroused long calls of

orang-utans differed only in temporal variables. Because Lameira & Wich (2008) were able to show that pulse duration is the most stable parameter over long distances, the arousal effect can still be perceived by distant audiences. Wich et al. (2003) also found temporal variables to be most important for context discrimination in Thomas langurs. A high arousal state also leads to an extended let-down phase in the form of bubbles, but their low amplitude makes this imperceptible at longer distances. Comparable results were obtained for 'pant hoots' in chimpanzee. Pant hoots emitted in contexts associated with increased arousal, locomotor activity, or vocal exertion more commonly included a let down-phase (Notman & Rendall 2005).

The only study to date focusing on context-dependent long calls was made by Galdikas & Insley (1988). They described a rare variant, the fast call. They characterized it as a close-range, post-contact, or conflict call (victory call). The fast call is classified as its own vocalization because both acoustic structure and context differ compared with the long calls analyzed in this study (Galdikas & Insley 1988; Hardus et al. 2009).

Long Call Functions

So far long calls were thought to function in both male spacing and female attraction (Mitani 1985b; Delgado 2003; Mitra Setia & van Schaik 2007). Our results suggest that the function of orang-utan long calls should be reconsidered according to long call context, at least for Bornean orang-utans. Spontaneously given long calls are supposed to serve in announcing a male's location (female attraction function). Tuanan females with dependent offspring move away from spontaneously emitted long calls, whereas on average sexually active females slightly approach long calls that support female attraction function as mate attraction in spontaneously emitted long calls (L. Dunkel unpubl. data). This contrasts to the situation in Sumatra, where females approach long calls of the dominant male irrespective of their reproductive state, and female attraction and female protection are therefore considered the primary functions of the long call (Mitra Setia & van Schaik 2007). Even though Mitra Setia & van Schaik (2007) did not differentiate between different contexts, it is very likely that spontaneously given long calls serve the female attraction function in Sumatra as well. Bornean females cannot reliably expect protection from flanged males since it is common that

they also force matings, as unflanged males do (Utami Atmoko et al. 2009). In Borneo, females with dependent offspring and male rivals would not be the primary target audience of spontaneously emitted long calls and could therefore be regarded as eavesdroppers. Costs for the caller can arise when there is a more dominant male eavesdropper.

Long calls in response to disturbance may serve to repel rival males (counter call, snag crashing) or potential predators (human observers). This is supported by the females' lack of ranging response to them compared with spontaneously emitted long calls, made possible by the females' ability to perceive contextual differences in long call acoustics even though these differences are graded and not discrete. Snag crashing is noisy display behavior and could therefore serve as long-distance signaling too, similar to chest beating in gorillas and drumming with hand and feet on chest, the ground, the tree trunks, and the buttresses of trees in chimpanzees (Schaller 1963; Arcadi et al. 1998). Conceivably, females might incorporate snag crashing as an additional signal in their information processing, but the noise of snag crashing may not reach as far as that of the calls, and in any case the minimal ranging responses after long calls given by disturbances from the environment (Fig. 5b) strongly suggest that females are able to differentiate contexts from long call acoustic structure alone. This low reaction to long calls of an aroused male could indicate that females have learned that males are somehow pre-occupied with repelling male rivals, obviating the need to avoid them. Thus, rival males may be the target audience of high-arousal long calls, whereas females or dominant male rivals may be eavesdroppers. Assuming that a male responds to another long calling male only when that caller is subordinate, costs could only arise when there is a third, more dominant male nearby.

We conclude that both individual and context differences are incorporated in long call acoustic structure and females perceive these differences, as shown in their differential reactions by long call context. However, the sample size of three individual flanged males in this study is rather small, and future work should evaluate the results obtained here with larger samples and including behavioral responses of different age- and sex-classes (unflanged males, flanged males, and sexually active females), as well as in other orangutan populations, especially in Sumatra.

Acknowledgements

Literature Cited

Arcadi, A. C., Robert, D. & Boesch, C. 1998: Buttress drumming by wild chimpanzees: temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness. Primates **39**, 505—518.

Crockford, C. & Boesch, C. 2003: Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. Anim. Behav. **66**, 115—125.

da Cunha, R. G. T. & Byrne, R. W. 2006: Roars of black howler monkeys (*Alouatta caraya*): evidence for a function in inter-group spacing. Behaviour **143**, 1169—1199.

Delgado, R. A. 2003: The Function of Adult Male Long Calls in Wild Orangutans (Pongo pygmaeus). Ph.D. Dissertation, Duke University, Durham.

Delgado, R. A. 2007: Geographic variation in the long calls of male orangutans (*Pongo* spp.). Ethology **113**, 487—498.

Delgado, R. A., Lameira, A. R., Ross, M. D., Husson, S. J., Morrogh-Bernard, H. C. & Wich, S. A. 2009: Geographic variation in orangutan long calls. In: Orangutans: Geographic Variation in Behavioral Ecology and Conservation (Wich, S. A., Utami Atmoko, S. S., Mitra Setia, T. & van Schaik, C. P., eds). Oxford University Press, Oxford, pp. 215—224.

Galdikas, B. F. M. 1983: The orangutan long call and snag crash at tanjung puting reserve. Primates **24**, 371—384.

Galdikas, B. M. F. 1985a: Adult male sociality and reproductive tactics among orangutans at Tanjung Puting. Folia Primatol. **45**, 9—24.

Galdikas, B. M. F. & Insley, S. J. 1988: The fast call of the adult male orangutan. J. Mammal. **69**, 371—375.

Hammerschmidt, K. & Todt, D. 1995: Individual-differences in vocalizations of young barbary macaques (*Macaca-sylvanus*) – a multi-parametric analysis to identify critical cues in acoustic signaling. Behaviour **132**, 381—399.

Hardus, M. E., Lameira, A. R., Singleton, I., Morrogh-Bernard, H. C., Knott, C. D., Ancrenaz, M., Utami Atmoko, S. S. & Wich, S. A. 2009: A description of the orangutan's vocal and sound repertoire, with a focus on geographic variation. In: Orangutans: Geographic Variation in Behavioral Ecology and Conservation (Wich, S. A., Mitra Setia, T. & van Schaik, C. P., eds). Oxford University Press, Oxford, pp. 49—60.

Hauser, M. D. 1996: Psychological design and communication. In: The Evolution of Communication. (Hauser, M. D., ed). MIT Press, Cambridge, MA. pp. 476—496.

Hirsig, R. 2002: Statistische Methoden in den Sozialwissenschaften: Eine Einführung im Hinblick auf Computergestützte Datenanalysen mit SPSS. Seismo Verlag, Zurich.

Lameira, A. R. & Wich, S. A. 2008: Orangutan long call degradation and individuality over distance: a playback approach. Int. J. Primatol. **29**, 615—625.

Mitani, J. C. 1985b: Sexual selection and adult male orangutan long calls. Anim. Behav. **33**, 272—283.

Mitani, J. C., GrosLouis, J. & Macedonia, J. M. 1996: Selection for acoustic individuality within the vocal repertoire of wild chimpanzees. Int. J. Primatol. **17**, 569—583.

Mitra Setia, T. & van Schaik, C. P. 2007: The response of adult orang-utans to flanged male long calls: inferences about their function. Folia Primatol. **78**, 215—226.

Norcross, J. L., Newman, J. D. & Cofrancesco, L. M. 1999: Context and sex differences exist in the acoustic structure of phee calls by newly-paired common marmosets (*Callithrix jacchus*). Am. J. Primatol. **49**, 165—181.

Notman, H. & Rendall, D. 2005: Contextual variation in chimpanzee pant hoots and its implications for referential communication. Anim. Behav. **70**, 177—190.

Rendall, D., Owren, M. J., Weerts, E. & Hienz, R. D. 2004: Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. J. Acoust. Soc. Am. **115**, 411—421.

Ross, M. D. & Geissmann, T. 2007: Call diversity of wild male orangutans: a phylogenetic approach. Am. J. Primatol. **69**, 305—324.

van Schaik, C. P. 1999: The socioecology of fission–fusion sociality in orangutans. Primates **40**, 69—86.

van Schaik, C. P., Wich, S. A., Utami, S. S. & Odom, K. 2005: A simple alternative to line transects of nests for estimating orangutan densities. Primates **46**, 249—254.

Schaller, G. B. 1963: The Mountain Gorilla: Ecology and Behavior. University of Chicago Press, Chicago.

Schehka, S., Esser, K. H. & Zimmermann, E. 2007: Acoustical expression of arousal in conflict situations in tree shrews (*Tupaia belangeri*). J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. **193**, 845—852.

Teixidor, P. & Byrne, R. W. 1999: The 'whinny' of spider monkeys: individual recognition before situational meaning. Behaviour **136**, 279—308.

Titus, K., Mosher, J. A. & Byron, K. W. 1984: Chance-corrected classification for use in discriminant analysis: ecological applications. Am. Midl. Nat. **111**, 1—7.

Utami Atmoko, S. S., Singleton, I., van Noordwijk, M. A., van Schaik, C. P. & Mitra Setia, T. 2009: Male-male relationships in orangutans. In: Orangutans Geographigc Variation in Behavioral Ecology and Conservation (Wich, S. A., Utami Atmoko, S. S., Mitra Setia, T. & van Schaik, C. P., eds). Oxford University Press, Oxford, pp. 235—244.

Waser, P. M. 1977: Individual recognition, intragroup cohesion and intergroup spacing-evidence from sound playback to forest monkeys. Behaviour **60**, 28—74.

Wich, S. A. & Nunn, C. L. 2002: Do male ''long-distance calls'' function in mate defense? A comparative study of long-distance calls in primates Behav. Ecol. Sociobiol. **54**, 474—484.

Wich, S. A., Koski, S., de Vries, H. & van Schaik, C. P. 2003: Individual and contextual variation in Thomas langur male loud calls. Ethology **109**, 1—13.

Zuberbühler, K., NoË, R. & Seyfarth, R. M. 1997: Diana monkey long-distance calls: messages for conspecifics and predators. Anim. Behav. **53**, 589—604.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Audio S1. Long call aroused individual.

Audio S2. Long call spontaneous individual.

Figure S1. Spectrogram view of pulse type 'roar' and newly recognized pulse type 'volcano.' Both pulses come from the same long call.

Figure S2. Differences in fundamental frequencies between individuals and contexts.

Figure S3. Differences in mean numbers of pulses between individuals and contexts.

Figure S4. Differences in mean call rate between individuals and contexts.

Table S1. Comparison of pulse type volcano and roar corrected for individuals ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

Table S2. Differences between individuals and contexts ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

Table S3. Context differences between individuals and within individuals ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.