

## Food-associated calling sequences in bonobos

Zanna Clay, Klaus Zuberbühler\*

School of Psychology, University of St Andrews

**Keywords:**

bonobo  
 call sequence  
 food-associated call  
*Pan paniscus*  
 primate vocal communication

When encountering food, chimpanzees, *Pan troglodytes*, and some other primates produce specific calls, whose acoustic structure covaries with the caller's food preference. For chimpanzees, there is evidence that listeners use these contingencies to guide their own foraging behaviour. Here, we investigated the vocal behaviour of another great ape, the bonobo, *Pan paniscus*, in response to food. We were able to distinguish acoustically five different call types given during interactions with food. These calls were not given singly, but usually as part of long and complex call sequences. We established the food preference hierarchies for 10 different individuals, housed at two different facilities. We found that the composition of call sequences produced by these individuals was not random, but was related to the type of food encountered by the caller. Significant variation in call composition was explained by taking into account the caller's individual food preferences, suggesting that bonobo food-calling sequences convey meaningful information to other group members.

Numerous mammals and birds produce calls in response to food and these signals commonly attract other group members (e.g. Dittus 1984; Elgar 1986a, b; Chapman & Lefebvre 1990; Heinrich & Marzluff 1991; Hauser & Marler 1993a, b; but see Gros-Louis 2004). A number of suggestions have been made to explain potential fitness benefits for this seemingly altruistic behaviour: decreasing predation risk because of increased vigilance (Elgar 1986a) or dilution (Pulliam & Caraco 1984; Newman & Caraco 1989; Ruxton 1995); benefiting kin (Hauser & Marler 1993a); attracting mates (Stokes & Williams 1971; Marler et al. 1986; Van Krunkelsven et al. 1996); attracting potential coalition partners (Van Krunkelsven et al. 1996); and announcing ownership to decrease risk of punishment from conspecifics (Hauser & Marler 1993b).

The proximate mechanisms and cognitive sophistication underlying the production of food calls have been subject to no less debate. In particular, the answer remains elusive as to whether food calls are simply inflexible and hardwired responses primarily driven by the arousal state of the signaller or more intentional and communicative acts that inform others about feeding events (e.g. Marler et al. 1992; Lieberman 1994). Whatever governs call production, various primate and bird studies have demonstrated that receivers can interpret food calls in terms of the event experienced by the caller, at least by having their attention referred to the event (e.g. toque macaques, *Macaca sinica*: Dittus 1984;

cottontop tamarins, *Saguinus oedipus*: Roush & Snowdon 2000; tufted capuchins, *Cebus apella*: Di Bitetti 2005). For example, experimental manipulations of food discovery and call production by Seabright fowl, *Gallus gallus*, have led to the interpretation of food calls as representational signals (Evans & Evans 2007). In some studies, call production was associated with food quantity or divisibility (e.g. chimpanzees, *Pan troglodytes*: Hauser & Wrangham 1987; spider monkeys, *Ateles geoffroyi*: Chapman & Lefebvre 1990; rhesus macaques, *Macaca mulatta*: Hauser & Marler 1993a; Hauser et al. 1993). In some other species, features of food quality appear to be encoded in the acoustic signal, mainly in terms of changes in call rates (domestic chickens, *Gallus domesticus*: Marler et al. 1986; Gyger & Marler 1988; cottontop tamarins: Elowson et al. 1991; Roush & Snowdon 2000; red-bellied tamarins, *Saguinus labiatus*: Caine et al. 1995; white-faced capuchins, *Cebus capucinus*: Boinski & Campbell 1996; Gros-Louis 2004) but also in terms of changes in call structure (e.g. golden lion tamarins, *Leontopithecus rosalia*: Benz et al. 1992; Benz 1993).

Perhaps one of the most complex systems has been described in rhesus macaques. These primates produce five acoustically distinct calls and production varies with the perceived food quality, although some call types are also produced in nonfood contexts (Hauser & Marler 1993a, b). Chimpanzees, in contrast, produce one main call type with significant acoustic gradation to food, the 'rough grunt' (Goodall 1965, 1968, 1986; Marler & Tenaza 1977). A series of studies has shown that the acoustic structure of this grunt vocalization covaries with perceived food quality, which is meaningful to other group members (Slocombe & Zuberbühler 2005, 2006).

\* Correspondence: K. Zuberbühler, School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JP, U.K.

E-mail address: kz3@st-andrews.ac.uk (K. Zuberbühler).

Relatively little is known about how bonobos, *Pan paniscus*, the closest relatives of chimpanzees, communicate about food. The bonobo vocal repertoire has been described in some detail (de Waal 1988; Bermejo & Omedes 1999) and its graded nature creates significant potential for encoding a wealth of world features experienced by the caller. In a study of wild bonobo vocalizations in Lilungu, Democratic Republic of Congo, Bermejo & Omedes (1999) classified the vocal repertoire into 15 graded basic call types. Rather than being given in isolation, individual calls were regularly combined to form longer vocal sequences. Although there was considerable variability, the authors were able to identify 19 principal sequences. These sequences were used across an array of behavioural contexts with some, such as feeding, eliciting a particularly diverse range of calls. The variable use of vocal sequences suggests a possibility for calls to be combined in different ways to provide different meanings. Another study of wild bonobo vocalizations in Lomako, Democratic Republic of Congo, suggested that bonobos also modified their vocalizations in response to social situations: some individuals were observed to adjust the acoustic structure of their long-distance vocalizations to match those of other group members (Hohmann & Fruth 1994).

Preliminary work on food-associated vocalizations in wild and captive settings has suggested that bonobos produce a range of call types when encountering food (de Waal 1988; Bermejo & Omedes 1999). This is unlike chimpanzees whose main food-associated call is the 'rough grunt' (Goodall 1965, 1968, 1986; Marler & Tenaza 1977). Furthermore, bonobos in both the wild and in captivity produce sequences of different food-associated vocalizations that frequently attract other group members to the same food source (de Waal 1988; Bermejo & Omedes 1999; Z. Clay, personal observation). This is particularly evident following the production of calls in response to highly preferred food items. In one experimental study, individuals called significantly more when discovering hidden food alone than when with others, regardless of food quantity (Van Krunkelsven et al. 1996). To what degree there was a strategic element to this pattern is difficult to assess, although the authors reported that vocalizing individuals obtained copulations and proximity to coalition partners.

In this study, we systematically examined the vocal behaviour of two groups of bonobos housed in San Diego, U.S.A. during controlled feeding events. Our aim was to describe the structure of the seemingly complex vocal behaviour previously reported and to examine whether patterns could be related to the callers' food preferences, and as such provide referential information to listening conspecifics.

## METHODS

### *Study Sites*

We collected data from two groups of captive bonobos at San Diego Zoo and San Diego Wild Animal Park between January and April 2008, with permission from the Animal Care Management at San Diego Zoo and Wild Animal Park, following approval by their IUCAC committee. The Zoo group consisted of three adult females, two adult males, one subadult female, one juvenile male and one juvenile female. The Park group consisted of three adult females, three adult males, one subadult female and one juvenile female. Both groups were housed in outdoor areas during the day (Zoo: 560 m<sup>2</sup>; Park: approximately 3000 m<sup>2</sup>) connected to heated indoor housing facilities, which served as sleeping areas at night (Zoo: one larger room, 136 m<sup>2</sup>, and four smaller rooms, each 55 m<sup>2</sup>; Park: one larger room, 47 m<sup>2</sup>, and three smaller rooms, each 40 m<sup>2</sup>). Individuals in both groups were unable to pass between the indoor and outdoor enclosures during the day. The Park group spent all

observation time together, while the composition of the Zoo group was managed to simulate a fission–fusion social system.

At both locations, individuals were usually fed in a group setting three to four times per day in both indoor and outdoor enclosures. At the Zoo, the diet consisted of 9% ape biscuits and cereals, 35% vegetables, 26% green leaf vegetables and 29% fruits. At the Park, the diet consisted of 17% ape biscuits and cereals, 20% vegetables, 24% green leaf vegetables and 39% fruits. At the Zoo, an artificial termite mound in the outdoor enclosure was filled daily with honey and human baby food for enrichment purposes. At both locations, individuals were fed a selection of approximately 25 different types of food per week. Individuals were also given separate supplementary enrichment feeds (such as ice lollies and seeds) several times per week. We collected data during routine feeds when food was scattered throughout the outdoor enclosure by a caregiver.

### *Food Preferences*

We tested food preferences of 10 adult individuals, five from each group. We conducted pairwise tests for 12 food types at the Zoo and 11 at the Park, following a procedure by Slocombe & Zuberbühler (2006). At the Zoo, individuals were allowed to choose two food items of similar size and shape from a tray. On two separate occasions each individual was presented with the same array repeatedly (range  $N = 2-3$  choices) with item location counterbalanced. At the Park, food preferences were established during regular lunchtime feeds, whereby subjects were individually presented with two different food items at least twice on two separate occasions. We recorded the individuals' first choices as a sign of preference and then calculated a preference score for each food item from all choices. To calculate a preference score, we counted the times each food type was chosen over the other food types and converted these 'first-choice' scores into percentages (high-preference foods = 67–100%, medium-preference foods = 34–66%, low-preference foods = 0–33%).

### *Vocal Behaviour*

We recorded vocalizations given by individuals while interacting with one type of food. We excluded any instances where individuals were either holding or eating more than one food type. All recordings were made at distances of 3–20 m with a Sennheiser MKH816 T directional microphone and Marantz PMD660 solid-state recorder that was set to a sampling rate of 44.1 kHz with 16 bits accuracy. If necessary, verbal comments were given and later transcribed. Audio recordings were transferred digitally onto a Toshiba laptop computer (Celeron 1.8 GHz), using Praat 4.3.17 ([www.Praat.org](http://www.Praat.org)).

### *Call Selection*

We recorded a total of 448 calling sequences from the 10 most vocally active adult individuals (Zoo: females LN, IK, LL; males YN, JU; Park: females: LT, LR; males EN, JJ, AK). The beginning of a call sequence was the point at which an individual made physical contact with a food item. We excluded any recordings that suffered from extensive background noise or other interference. We then randomly selected, for each individual, three calling sequences from the high-, medium- and low-preference classes ( $N = 90$ ). Within each preference class, calls were selected randomly with regard to food type. Because sequences varied in the number of calls produced (approximate range 1–40) we only considered calls within the first three calls of a sequence of at least three calls ( $N = 270$  calls).

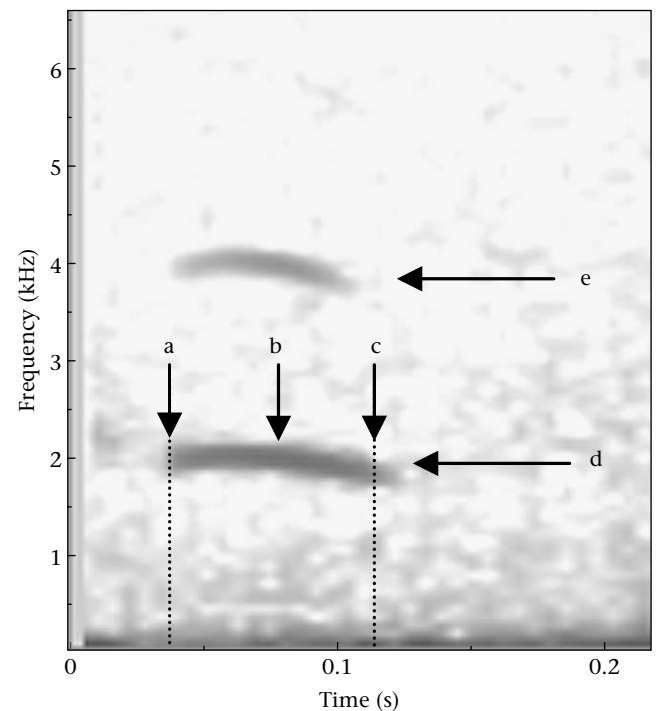
## Acoustic Analyses

We carried out quantitative analyses of the acoustic structure of the different vocalizations, using Praat 4.3.17. One call type, the grunts, had a fundamentally different acoustic structure to the other food-associated vocalizations, which required a different set of acoustic parameters. Grunts were much noisier, typically unvoiced, and with strong formant bands. For these, we used the following settings: analysis window length 0.025 s, dynamic range 30 dB and spectrogram window length 0.005 s. For all other calls, we used the following settings: pitch range 500–2500 Hz, optimized for voice analysis, spectrogram view range 0–20 kHz (to determine the number of harmonics) and 0–5 kHz (window length 0.01 s, dynamic range 70 dB) to measure fundamental frequency. We performed pitch analysis, using a script written by M. Owren (personal communication), and verified the generated values using the harmonic cursor. All further spectral measurements were taken from the fundamental frequency (F0): (1) mean fundamental frequency (Hz): average F0 across the entire call; (2) transition onset ( $\Delta$ Hz): frequency of maximum energy at call onset minus frequency of maximum energy at call middle; (3) transition offset ( $\Delta$ Hz): frequency of maximum energy at call middle minus frequency of maximum energy at call offset; (4) overall transition ( $\Delta$ Hz): frequency of maximum energy at call end minus frequency of maximum energy at call onset; (5) maximum fundamental frequency (Hz): maximum frequency of F0; (6) minimum fundamental frequency (Hz): minimum frequency of F0; (7) peak time: location in the temporal domain where maximum acoustic energy occurs, expressed as a proportion of the call duration; (8) number of harmonics: number of harmonic bands visible. As grunts were mainly unvoiced, they did not possess an F0 produced by oscillations of the vocal folds. We thus counted the oscillations visible in the spectrogram produced by other sound sources, divided by duration of the call. In the temporal domain, we measured for all calls the call duration (first three calls only). Figure 1 illustrates the various acoustic parameters.

We screened the data for outliers by producing standardized Z scores. We rejected calls with a Z score greater than 3.29 in one or more parameters (Tabachnick & Fidell 2001). We then regressed all parameters to check for collinearity and removed parameters with a variance inflation factor greater than 10. We then conducted a discriminant function analysis (DFA) to assess whether each of the uncorrelated acoustic variables, when combined in one model, could discriminate between the four graded call types. Because DFA is sensitive to unbalanced data, each of the 10 individuals equally contributed five randomly selected calls for each of the four vocalizations ( $N = 200$  calls, excluding grunts). Therefore, 50 calls were entered for each vocalization.

We used the leave-one-out classification procedure to cross-validate the discriminant function that was generated. In this procedure, each call is classified by the functions derived from all calls other than that one. Since the acoustic data for food-associated calls were two-factorial (caller ID; call type), it has been argued that conventional DFA does not allow for a valid estimation of the overall significance of discriminability (Mundry & Sommer 2007). Therefore, we used a permuted discriminant function analysis (pDFA), using a macro written by R. Mundry and C. Sommer (Mundry & Sommer 2007), to estimate the significance of the number of correctly classified calls (cross-validated).

We ran one-way related-samples analysis of variance tests to examine whether each of the acoustic parameters varied statistically with each call type. Grunts were included in the analyses of call duration and mean fundamental frequency. Each individual ( $N = 10$ ) contributed a mean value per call type per parameter, which was derived from five calls per call type category ( $N = 250$



**Figure 1.** Temporal and structural parameters measured on food-associated calls: call duration (s) =  $c - a$ ; fundamental frequency, F0 (Hz) =  $d$ ;  $N$  harmonics (1 in this call) =  $e$ ; transition onset ( $\Delta$ Hz) = frequency of maximum energy at call onset (a) – frequency of maximum energy at call middle (b); transition offset ( $\Delta$ Hz) = frequency of maximum energy at call middle (b) – frequency of maximum energy at call offset; overall transition ( $\Delta$ Hz) = frequency of maximum energy at call end (c) – frequency of maximum energy at call onset (a). Depicted is a time–frequency spectrogram of a peep vocalization made by adult female LR.

raw calls). We conducted post hoc pairwise Sidak-corrected comparisons to examine whether any of the acoustic parameters could discriminate between the call types.

In addition, we carried out an interobserver reliability test for call type classification using two naïve observers. After completing a training set of preclassified calls (randomly selected 10%, of original call set), the naïve observers independently classified 10% of the original call set ( $N = 30$ ). The test set comprised an equal selection of each of the five call types, all randomly selected. We calculated Cohen's Kappa coefficients to determine whether the levels of observer agreement reached the standard accepted level (Cohen's  $\kappa = 0.80$ ).

## Structure of Call Sequences

The focus of the second major analysis concerned the structure and composition of the call sequences. We analysed 236 raw call sequences produced by 10 individuals. Each individual contributed a range of two to five call sequences per food type. Means of the raw call sequences were then calculated for three randomly selected different food types for each of the three preference categories. We measured (1) the absolute number of each call type (first 30 s per sequence) and (2) the relative proportion of different call types (first 30 s per sequence), (3) the intercall interval (first three calls only), and (4) the call rate ( $N$  calls within first 30 s per sequence). Following statistical validation of call type categorization in the first stage of acoustic analysis and interobserver reliability tests, call types in this analysis were those assigned by human raters.

We calculated the mean number and relative proportion of each call type produced in sequence to high-, medium- and low-preference foods. One overall mean per combination of individual and food preference category was entered. For each individual, the overall mean was calculated from the means of three randomly selected food types for each of the three food preference levels. For both analyses, we used a matched-pairs design using Friedman and Wilcoxon signed-ranks tests (exact, two-tailed) and a Sidak correction to minimize the risk of family-wise errors.

For the intercall interval analysis, we used two-way analysis of variance tests with food preference entered as the fixed factor (high, medium, low) and caller identity as the random factor. Data were analysed from all 10 individuals, with intercall intervals taken from three randomly selected food types for each of the three preference categories ( $N = 90$  sequences). We calculated the median of the first three intercall intervals within each sequence.

For the call rate analysis, we calculated the mean number of calls produced within the first 30 of a sequence, where each individual contributed a mean call rate per preference class, taken from the means of three randomly selected food types. As the call rate data were not normally distributed we conducted nonparametric Friedman and Wilcoxon signed-ranks tests (exact, two-tailed).

All statistical tests were carried out using SPSS version 12.0 (SPSS Inc., Cary, NC, U.S.A.) except for the pDFA which was conducted using R 2.8.1. (The R Foundation for Statistical Computing, Vienna, Austria). All tests were two tailed and alpha levels were set at 0.05, unless stated as being corrected.

## RESULTS

### Food Preferences

We conducted pairwise choice tests for all possible combinations of 12 food types at the Zoo and 11 food types at the Park. While there was some consistency of food preferences across individuals, particularly for the most preferred foods, we found some minor individual differences. Nevertheless, we were able to assign the different foods into three broad preference classes for each individual, based on the preference scores: high (67–100% first choices); medium (34–66% first choices); low (0–33% first choices; Table 1).

**Table 1**  
Results of food preference tests conducted on 10 adult bonobos at the San Diego Zoo and Wild Animal Park

Food	Zoo individuals					Park individuals				
	LN	LL	IK	YN	JU	LT	LR	EN	AK	JJ
Fig	—	—	—	—	—	100	100	100	100	100
Raisin	100	100	100	95	91	90	90	90	90	90
Grape	91	91	91	95	91	70	75	80	75	80
Banana	77	77	73	82	91	—	—	—	—	—
Popcorn	73	73	<b>55</b>	<b>60</b>	<b>60</b>	—	—	—	—	—
Apple	<b>55</b>	<b>64</b>	<b>59</b>	<b>59</b>	<b>55</b>	<b>60</b>	<b>60</b>	<b>65</b>	70	70
Orange	<b>65</b>	<b>55</b>	<b>66</b>	<b>64</b>	<b>64</b>	<b>50</b>	<b>50</b>	<b>50</b>	<b>45</b>	<b>50</b>
Biscuit	<b>46</b>	<b>36</b>	32	32	<b>36</b>	<b>65</b>	<b>65</b>	<b>65</b>	<b>60</b>	<b>60</b>
Celery	—	—	—	—	—	25	<b>40</b>	<b>40</b>	<b>40</b>	<b>40</b>
Melon	27	<b>34</b>	14	<b>36</b>	<b>41</b>	—	—	—	—	—
Lettuce	18	18	14	18	18	25	30	25	30	30
Yam	9	18	18	18	9	8	20	10	10	10
Pepper	9	31	9	27	18	10	0	0	0	0

Preference scores (shown as percentages) were based on the relative number of times a food was chosen over all foods tested. The foods are ordered by rank preference averaged across individuals: 0–30, low preference; 34–66, medium preference (bold); 67–100, high preference (italic). Dashes indicate foods that were not tested.

### Vocal Behaviour

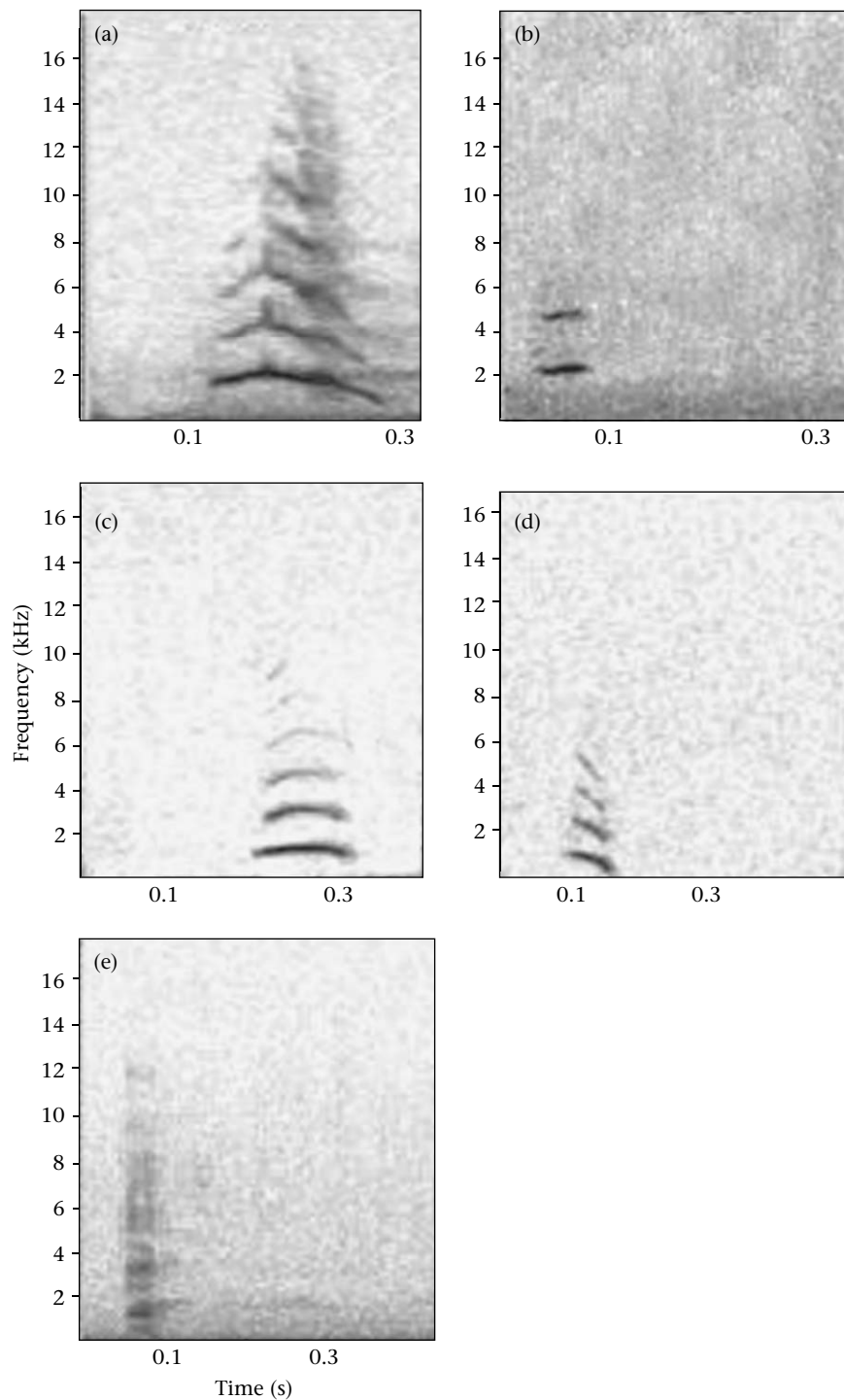
#### Acoustic analyses

Our preliminary observations suggested that in response to food, bonobos produced five perceptually distinct call types: barks, peeps, peep-yelps, yelps and grunts (Fig. 2).

After checks for multicollinearity and singularity, we were able to enter seven uncorrelated of the nine original acoustic parameters for further analyses ( $N = 200$  calls, excluding grunts): call duration, peak time, mean fundamental frequency, number of harmonics, transition onset, transition offset and overall transition. We conducted a DFA, derived from all seven acoustic variables, to assess how well each of the acoustic variables could discriminate between the four graded call types (bark, peep, peep-yelp and yelp). Of the three functions used in the DFA, two functions significantly discriminated between the call types (see Fig. 3). The functions explained a significant amount of the variation in the acoustic structure of the call types (Wilk's lambda = 0.089,  $\chi^2_{21} = 468.718$ ,  $P < 0.001$ ). In a cross-validated analysis, the functions successfully classified 86% (172/200; cross-validated) of the calls according to call type, a level well above that expected by chance. The success rate of classification of call types was highest for barks and peeps, followed by yelps and then peep-yelps (correct classification for barks = 96%, peeps = 94%, yelps = 80%, peep-yelps = 74%). We then used a pDFA (Mundry & Sommer 2007) to estimate the significance of the number of correctly classified calls (cross-validated). Results from the pDFA indicated a highly significant level of discrimination ( $P = 0.001$ ).

In addition, to verify whether our original call classification was reliable across human raters, we conducted an interobserver reliability test with two naïve observers. The observers classified a random selection of 10% of the original call set, and observers independently attained 97% and 93% agreement with our original classification. Our scores showed very high levels of agreement (for observers 1 and 2, respectively: Cohen's  $\kappa = 0.96, 0.92$ ).

To examine whether each of the uncorrelated acoustic parameters varied statistically between call types, we conducted one-way related-samples analysis of variance tests with call type as the fixed factor (sphericity assumed). Grunts were included in the analyses of call duration and mean fundamental frequency. Both parameters varied statistically among call types (mean F0:  $F_{4,36} = 329.409$ ,  $P < 0.001$ ; call duration:  $F_{4,36} = 10.300$ ,  $P < 0.001$ ). We also found that the five remaining acoustic parameters varied consistently among the four nongrunt call types ( $N$  harmonics:  $F_{3,27} = 30.071$ ,  $P < 0.001$ ; peak time:  $F_{3,27} = 6.299$ ,  $P = 0.033$ ; transition onset:  $F_{3,27} = 33.080$ ,  $P < 0.001$ ; transition offset:  $F_{3,27} = 10.894$ ,  $P < 0.001$ ). One acoustic parameter, overall transition, failed to reach significance ( $F_{3,27} = 2.908$ ,  $P = 0.053$ ). Post hoc pairwise Sidak-corrected comparisons revealed that mean fundamental frequency significantly discriminated among all call types except for the two highest-pitched call types, the barks and peeps. Barks were significantly longer in duration and had more harmonic bands visible than any other call type. Peeps were the shortest call type and showed significantly fewer harmonics than other call types. Barks had a pointed acoustic shape with steep upward (transition onset) and downward (transition offset) strokes in acoustic energy. Their steep upward stroke distinguished them from peeps, peep-yelps and yelps and their downward stroke distinguished them significantly from peeps and yelps. Yelps showed a distinctive overall downward stroke form, which discriminated them significantly from the acoustically flat peeps in both the onset and offset transition, and from the upward curving peep-yelps in the onset transition. Barks also had a significantly later peak time than peeps or yelps. In contrast to the other graded call types, grunts showed a markedly different structure; they were distinctly noisy, much



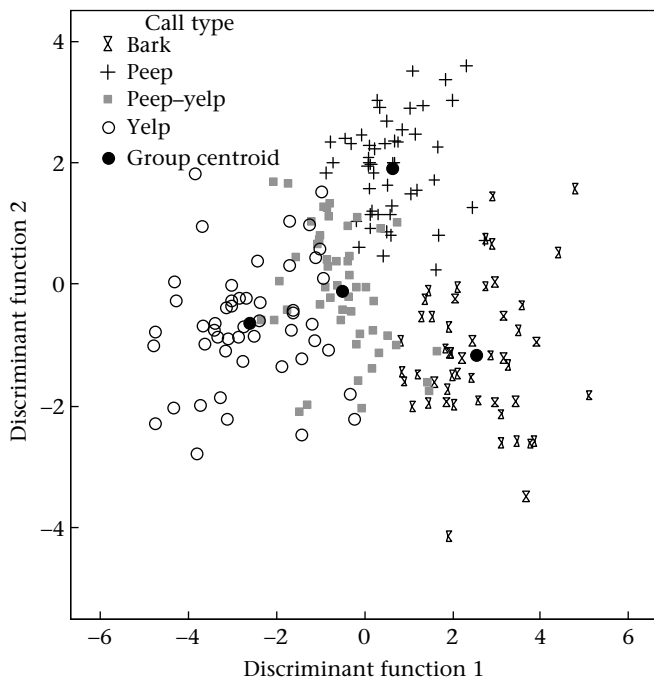
**Figure 2.** Examples of time–frequency spectrograms illustrating five different types of food-associated calls, produced by EN, an adult male bonobo: (a) bark, (b) peep, (c) peep-yelp, (d) yelp, (e) grunt.

lower in pitch, mostly unvoiced, with clearly visible formants and no visible fundamental frequency and harmonic bands. Table 2 and Fig. 4 summarize the results.

#### Structure of Call Sequences

To describe the different call sequences produced in response to food, we first compared the absolute number of each call type per sequence. Each individual contributed an overall mean per preference class (calculated from the means of three food types, i.e.  $N = 9$

total for 10 individuals). We conducted Friedman tests and found that the absolute frequency of all five call types varied significantly among preference classes (barks:  $\chi^2_2 = 15.077$ ,  $P < 0.001$ ; peeps:  $\chi^2_2 = 16.632$ ,  $P < 0.001$ ; peep-yelps:  $\chi^2_2 = 8.6$ ,  $P = 0.012$ ; yelps:  $\chi^2_2 = 15.436$ ,  $P < 0.001$ ; grunts:  $\chi^2_2 = 7.913$ ,  $P = 0.017$ ). With a corrected alpha level of 0.0169, post hoc Wilcoxon signed-ranks tests revealed that significantly more barks and peeps occurred in sequences associated with high- than low-preference foods (barks:  $Z = -2.521$ ,  $N = 10$ ,  $P = 0.008$ ; peeps:  $Z = -2.803$ ,  $N = 10$ ,  $P = 0.002$ ) or medium-preference foods (barks:  $Z = -2.521$ ,  $N = 10$ ,



**Figure 3.** Distribution of discriminant scores along the two canonical discriminant functions established to discriminate different bonobo food-associated calls (Eigen values: Function 1=3.275; Function 2=1.351). Black circles represent group centroids. The four graded call types from the original classification overlies the discriminant function scores.

$P = 0.008$ ; peeps:  $Z = -2.803$ ,  $N = 10$ ,  $P = 0.002$ ). Significantly more peep-yelps occurred in sequences associated with medium- than low-preference foods ( $Z = -2.803$ ,  $N = 10$ ,  $P = 0.002$ ) and significantly more yelps occurred in sequences associated with

**Table 2**  
Results of post hoc Sidak-corrected comparison tests for differences between sample means of acoustic parameters of five food-associated call types

Acoustic parameter	Call type	Bark	Peep	Peep-yelp	Yelp	Grunt
Call duration	Bark		**	*	*	**
	Peep					
	Peep-yelp					
	Yelp					
	Grunt					
Mean fundamental frequency	Bark			***	***	***
	Peep			***	***	***
	Peep-yelp				***	***
	Yelp					***
	Grunt					
Peak time	Bark		*		**	
	Peep					
	Peep-yelp					
	Yelp					
Number of harmonics	Bark		***	**	*	
	Peep			**	**	
	Peep-yelp					
	Yelp					
Transition onset	Bark		***	**	***	
	Peep				*	
	Peep-yelp				*	
	Yelp					
Transition offset	Bark		**		**	
	Peep				**	
	Peep-yelp					
	Yelp					

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

low- and medium- than high-preference foods (both:  $Z = -2.805$ ,  $N = 10$ ,  $P = 0.002$ ). Finally, we found trends of increased production of grunts in sequences associated with low- and medium- than high-preference foods (high to low:  $Z = -2.201$ ,  $N = 10$ ,  $P = 0.031$ ; high to medium:  $Z = -2.023$ ,  $N = 10$ ;  $P = 0.063$ ). Figure 5 summarizes the results.

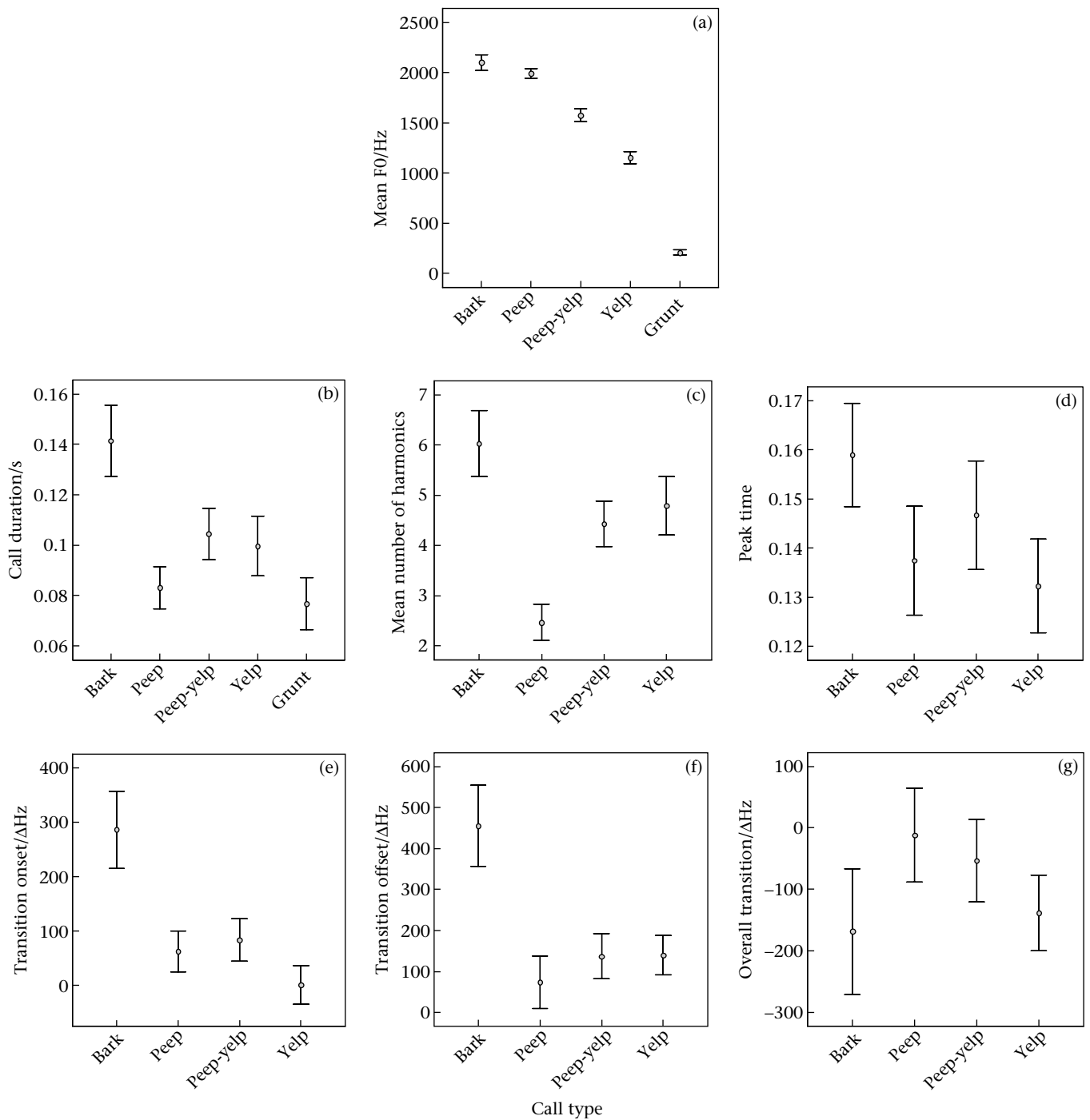
We also conducted a Friedman test to examine the relative contributions of call types to call sequences. We found that the proportion of barks ( $\chi^2_2 = 15.44$ ,  $P < 0.001$ ), peeps ( $\chi^2_2 = 17.90$ ,  $P < 0.001$ ) and yelps ( $\chi^2_2 = 14.00$ ,  $P < 0.001$ ) varied significantly among preference classes. Proportions of grunts also tended to be associated with preference classes ( $\chi^2_2 = 5.48$ ,  $P = 0.061$ ), but no significant differences were found for peep-yelps. With a corrected alpha level of 0.0169, post hoc Wilcoxon signed-ranks tests revealed significantly higher proportions of both peeps and barks associated with high- than low-preference foods (peeps:  $Z = -2.805$ ,  $N = 10$ ,  $P = 0.002$ ; barks:  $Z = -2.521$ ,  $N = 10$ ,  $P = 0.008$ ) or medium-preference foods (peeps:  $Z = -2.803$ ,  $N = 10$ ,  $P = 0.002$ ; barks:  $Z = -2.521$ ,  $N = 10$ ,  $P = 0.008$ ). Significantly higher proportions of yelps occurred in sequences associated with medium- than high-preference foods ( $Z = -2.808$ ,  $N = 10$ ,  $P = 0.002$ ) and low- than to high-preference foods ( $Z = -2.666$ ,  $N = 10$ ,  $P = 0.004$ ). Finally, there was a trend of higher proportions of grunts occurring in sequences associated with low- and medium- than high-preference foods ( $Z = -2.666$ ,  $N = 10$ ,  $P = 0.046$ ;  $Z = -2.808$ ,  $N = 10$ ,  $P = 0.043$ , respectively). Figure 6 summarizes the results.

While we found no significant effect of preference class on intercall interval ( $F_{2,60} = 3.024$ ,  $P = 0.073$ ), there was a significant effect of call rate ( $\chi^2_2 = 7.2$ ,  $N = 10$ ,  $P = 0.03$ ). With a corrected alpha of 0.0169, post hoc analyses using Wilcoxon signed-ranks tests revealed a trend for more calls to be produced in association with high- than medium- or low-preference foods (high versus medium:  $Z = -2.293$ ,  $N = 10$ ,  $P = 0.02$ ; high versus low:  $Z = -2.090$ ,  $N = 10$ ,  $P = 0.037$ ), but there was no difference in number of calls produced in association with foods of medium preference compared to low preference.

## DISCUSSION

Our study demonstrated empirically that captive bonobos at two locations produced five acoustically distinct call types when interacting with food: barks, peeps, peep-yelps, yelps and grunts. The production and distribution of these call types within a sequence was not random but was significantly associated with the preference score of the food. Barks were produced almost exclusively in association with highly preferred foods, and peeps were given in significantly higher proportions to high- than to medium- or low-preference food. Peep-yelps, the intermediate call type, were produced irrespective of food preference. Significantly higher proportions of yelps were produced in calling sequences associated with medium- and low- than to high-preference foods. While we observed similar trends for grunts and yelps, these effects were not significant. Furthermore, there were trends of increased production of yelps and grunts with decreasing food preference but statistical discrimination between medium- and low-preference foods was not possible. The general lack of strong distinctions between medium- and low-preference foods may be caused by an insensitivity of the acoustic parameters chosen or because bonobos genuinely only make strong vocal distinctions between high- and nonhigh-preference foods.

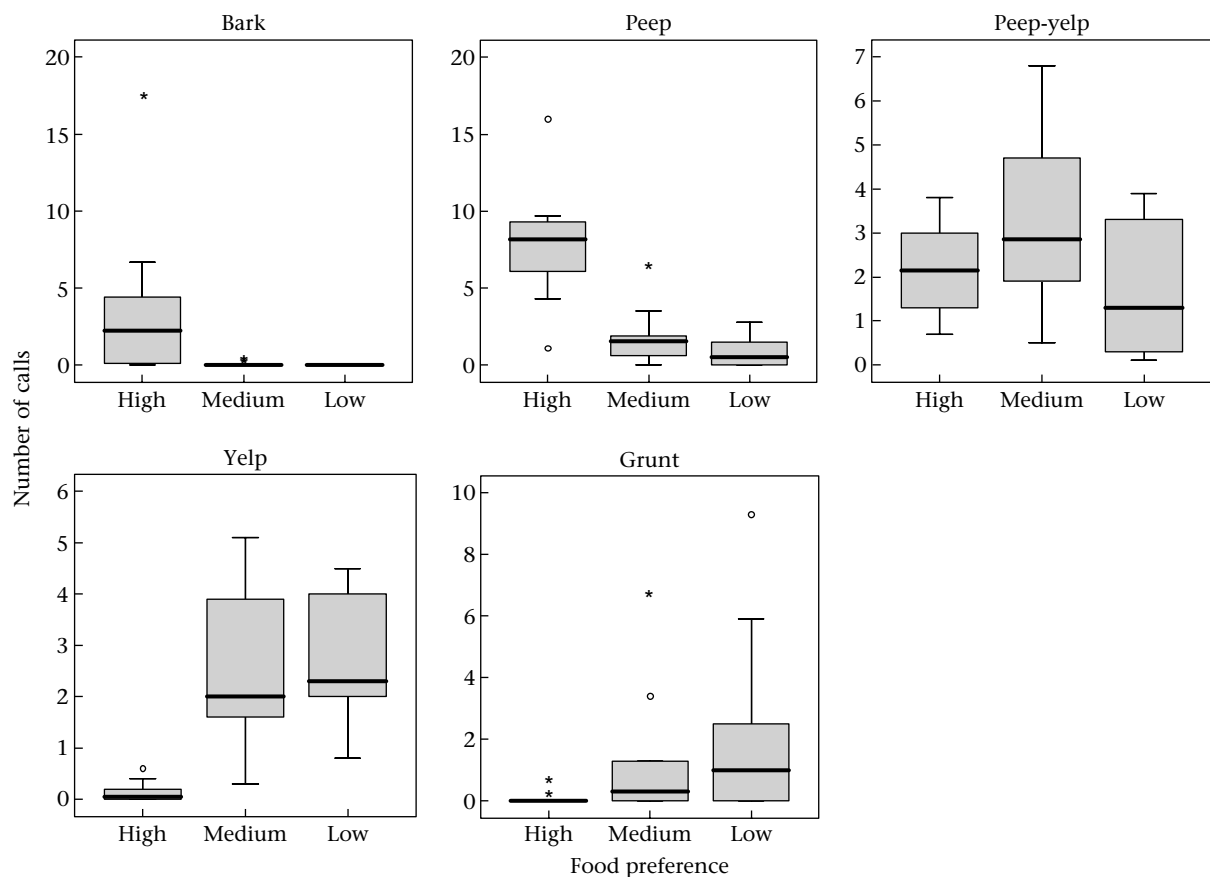
The five food-associated calls produced by the bonobos in San Diego lay on a graded pitch continuum, with barks at the high end, followed by peeps, peep-yelps, yelps and finally grunts. The grunts we observed were the most acoustically distinct call type, owing to



**Figure 4.** Mean values  $\pm$  SE for seven acoustic parameters showing the similarities and differences between the different food-associated call types produced by bonobos at San Diego Zoo and Wild Animal Park: (a) fundamental frequency, F0 (Hz); (b) call duration (s); (c) number of harmonics; (d) peak time; (e) transition onset ( $\Delta$ Hz); (f) transition offset ( $\Delta$ Hz); (g) overall transition ( $\Delta$ Hz). The last five parameters are missing for grunts because the calculation of the F0 required to calculate these parameters was not possible. This was because grunts were typically unvoiced and so did not possess an F0 distinguishable in the spectrogram.

their atonal or noisy nature and lack of energy in the fundamental frequency band or harmonics. Barks were longest in duration, characterized by a distinctive pointed shape and numerous visible harmonic bands. While peeps were also high pitched, they were shorter than barks and of simpler structure with only few harmonic bands and a flat frequency contour. Although yelps and peep-yelps were acoustically similar, peeps were noticeably shorter in duration while yelps possessed a marked downward stroke frequency contour, in contrast to the arched contour of the lower-pitched peep-yelps.

Analysis of the intercall interval between calls indicated that food preference did not affect the speed by which calls were produced. However, more calls were produced in sequences associated with high-preference foods, suggesting that although rate did not increase, calls were produced in longer sequences. In a study on rhesus macaques, call rates were explained by differences in the callers' hunger levels (Hauser & Marler 1993a, b). However, hunger levels were unlikely to play a role in this study as the bonobos were fed the same quantity of food at the same time of day when the calls were recorded.



**Figure 5.** Box plots showing the number of times each of the five call types was produced within food-associated call sequences. Thick black lines represent medians; box edges represent the upper and lower hinges of the H-spread, which generally matches the upper and lower quartiles; whiskers represent the adjacent values, which are the most extreme values still lying within hinges and the normal distribution of the sample. Outliers are marked with circles and extreme cases with asterisks.

#### *Food-associated Calls in Chimpanzees and Bonobos*

A principal motivation of this study was to compare food-associated calls produced by bonobos with those of their closest relatives, the chimpanzees. Earlier studies have found that chimpanzees produce one main type of food-associated call, the rough grunt (Goodall 1986). This call type is highly graded and changes in the acoustic morphology are related to perceived food quality (Slocombe & Zuberbühler 2006), which is used by listeners to draw inferences about the food discovered by the caller (Slocombe & Zuberbühler 2005). While the bonobos in this study also produced grunts, we found that their grunts were mainly associated with low-preference foods and that this call type represented only a small fraction of all calls produced upon contact with food. An ongoing study of individuals at Lola Ya Bonobo Sanctuary, Democratic Republic of Congo, suggests that there may be greater vocal diversity within the grunt call itself than was observed in the bonobos at San Diego, and it may be used more widely by certain individuals (Z. Clay, unpublished data). Nevertheless, in contrast to chimpanzees, bonobos appear to communicate information about perceived food quality by producing a range of call types, while varying the production and distribution of each type in a calling sequence. The evolution of this species difference is difficult to understand, but it demonstrates that vocal behaviour in the food context can diverge relatively rapidly.

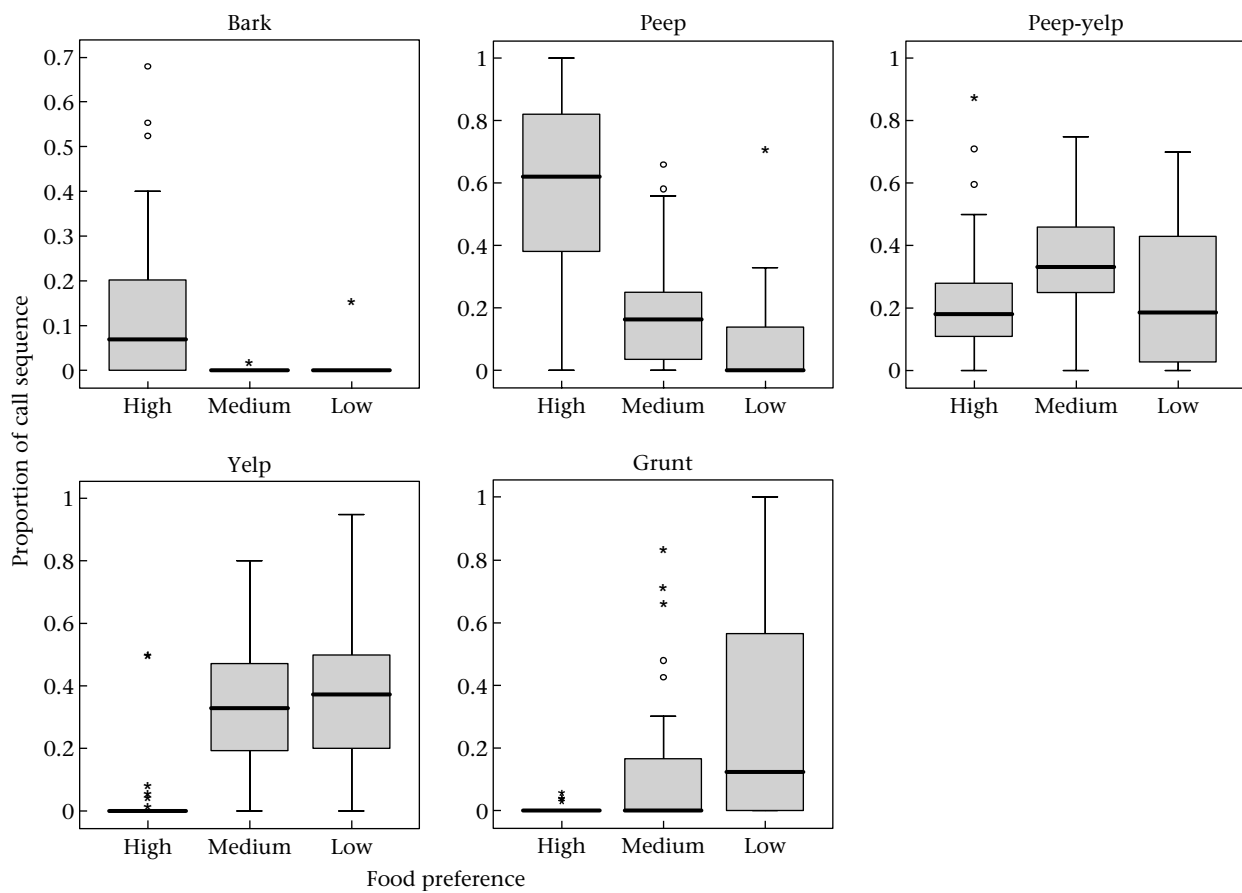
The vocal behaviour observed in bonobos in this study resembled the pattern described in rhesus macaques in some ways (Hauser & Marler 1993a). Rhesus macaques produce five call types in response to foods, some of which are associated with highly

preferred or rare foods (warble, harmonic arches and chirps). Coos and grunts are produced in both food and nonfood contexts, and only grunts differ between these contexts. In our study, we found that only the production of peep-yelps occurred indiscriminately across preference classes, similar to the rhesus monkeys' coo calls. Peep-yelps were also produced in a range of other nonfood contexts, and thus more systematic work examining acoustic morphology is required before we can make more precise comparisons. In significant contrast to macaques, four of five food-related calls given by bonobos are part of an acoustic continuum (Fig. 2).

#### *Function and Meaning of Bonobo Food Calls*

Our results suggest that bonobos can communicate meaningfully about an important type of external event, the discovery of food. The mechanism underlying call production (for example, whether the result of a deliberate attempt to inform others or a mere reflection of changes in arousal) largely remains obscure. Barks are the highest-pitched call type, reliably given to highly preferred foods, and it might be argued that this is in line with an arousal-based explanation (Marler 1977; Marler et al. 1992). As always, arousal remains an elusive concept, and finding independent measures to quantify it has proved extremely challenging, as illustrated well by this study.

Another important observation is that some of the calls described, particularly the peeps and peep-yelps, appear also to be produced in nonfood contexts, such as during mother-infant interactions, grooming, alarm, travel and also after agonistic



**Figure 6.** Box plots showing the relative proportions of the five call types present in food-associated calling sequences. See Fig. 5 for an explanation of the box plots.

encounters (de Waal 1988; Bermejo & Omedes 1999; Z. Clay, personal observations). Of course it is always possible that there is subtle but consistent context-specific acoustic variation within the call types, a topic for future research. Alternatively, it is possible that bonobos use these calls in a range of contexts, suggesting that their communicative significance is broader than referring a listener's attention to the fact that food has been discovered. Although we are unable to estimate accurately the referential specificity of these calls, the fact that overall sequence composition varies reliably between food preference classes suggests that receivers can extract accurate inferences about the nature of an ongoing feeding event, by paying attention to the structure of the sequence.

Vocal advertisement of feeding patches may present an evolutionary conundrum. From the signaller's perspective, such behaviour is costly if it leads to loss of food to competitors (Elgar 1986a). Callers could minimize these costs if they took into account, for example, food patch size, divisibility or composition of the nearby audience (Hauser & Wrangham 1987; Hauser et al. 1993; Zuberbühler 2008). It has also been argued that immediate costs may be outweighed by long-term benefits, for instance if callers receive social benefits, such as augmenting ties with allies (Van Krunkelsven et al. 1996), preventing punishment from dominant individuals (Hauser & Marler 1993b), or attracting potential mates (Stokes & Williams 1971; Marler et al. 1986; Van Krunkelsven et al. 1996). Studies on audience composition are required to obtain a better understanding of the impact of such social factors on call production (e.g. Slocombe & Zuberbühler 2007).

The fact that bonobos produce reliable signals to low-quality foods is more difficult to explain. Individuals may differ in their

day-to-day needs and, as such, any information about food discoveries can be important to receivers to decide whether to abandon a current activity. In their natural African forests, primates are confronted with complex and continuously changing food availability (Janmaat et al. 2006a, b). Visual information is difficult to obtain in dense forests, rendering vocal signals an extremely efficient way to update knowledge.

While many avenues are still open to explore, our study has shown that bonobos alter the production of different types of vocal signals as a function of perceived quality of food. Although some of these calls appear to lack high referential specificity, others are only given to very specific preference classes. More importantly, the combination of the different call types into sequences generates a reliable vocal signal that is useful for nearby listeners to make inferences about an external event. Playback studies will ultimately be required to determine how exactly the sequence patterns described in this study convey meaning to receivers and whether they use information conveyed in call sequences to guide their foraging decisions. If food-associated call sequences provide information about the food being eaten, we should expect receivers to modify their foraging behaviour based on what they had heard and navigate more effectively to food associated with the call.

#### Acknowledgments

We are very grateful to the San Diego Zoo and Wild Animal Park for enabling us to carry out this study. We are also especially grateful to Thibaud Gruber for his help with data collection. We thank all the Zoo and Park bonobo caretakers and staff, especially the head keepers Mike Bates and Donna Lundy and also Kim

Livingstone and Michelle Stancer for their assistance. We thank Eric Bowman and Mike Oram for statistical advice and Katie Slocombe, Simon Townsend and three anonymous referees for their advice. The study has been funded by a grant from the EC-Pathfinder FP-6 initiative ('What it means to be human') and a Research Leadership Award by the Leverhulme Trust.

## References

- Benz, J. J.** 1993. Food-elicited vocalizations in golden lion tamarins: design-features for representational communication. *Animal Behaviour*, **45**, 443–455.
- Benz, J. J., Leger, D. W. & French, J. A.** 1992. Relation between food preference and food-elicited vocalizations in golden lion tamarins (*Leontopithecus rosalia*). *Journal of Comparative Psychology*, **106**, 142–149.
- Bermejo, M. & Omedes, A.** 1999. Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, **70**, 328–357.
- Boinski, S. & Campbell, A. F.** 1996. The 'huh' vocalization of white-faced capuchins: a spacing call disguised as a food call. *Ethology*, **102**, 826–840.
- Caine, N. G., Addington, R. L. & Windfelder, T. L.** 1995. Factors affecting the rates of food calls given by red-bellied tamarins. *Animal Behaviour*, **50**, 53–60.
- Chapman, C. A. & Lefebvre, L.** 1990. Manipulating foraging group size: spider monkey food calls at fruiting trees. *Animal Behaviour*, **39**, 891–896.
- Di Bitetti, M. S.** 2005. Food-associated calls and audience effects in tufted capuchin monkeys, *Cebus apella nigrilus*. *Animal Behaviour*, **69**, 911–919.
- Dittus, W. P. J.** 1984. Toque macaque food calls: semantic communication concerning food distribution in the environment. *Animal Behaviour*, **32**, 470–477.
- Elgar, M. A.** 1986a. The establishment of foraging flocks in house sparrows: risk of predation and daily temperature. *Behavioral Ecology and Sociobiology*, **19**, 433–438.
- Elgar, M. A.** 1986b. House sparrows establish foraging flocks by giving chirrup calls if the resources are divisible. *Animal Behaviour*, **34**, 169–174.
- Elowson, A. M., Tannenbaum, P. L. & Snowdon, C. T.** 1991. Food-associated calls correlate with food preferences in cotton-top tamarins. *Animal Behaviour*, **42**, 931–937.
- Evans, C. S. & Evans, L.** 2007. Representational signalling in birds. *Biology Letters*, **3**, 8–11.
- Goodall, J.** 1965. Chimpanzees of the Gombe Stream Reserve. In: *Primate Behavior: Field Studies of Monkeys and Apes* (Ed. by I. DeVore), pp. 425–473. New York: Holt, Rinehart & Winston.
- Goodall, J.** 1968. A preliminary report on expressive movements and communication in the Gombe Stream chimpanzees. In: *Primate Studies in Adaptation and Variability* (Ed. by P. C. Jay), pp. 313–374. New York: Holt, Rinehart & Winston.
- Goodall, J.** 1986. *The Chimpanzees of Gombe: Patterns of Behavior*. Boston, Massachusetts: Belknap Press of the Harvard University Press.
- Gros-Louis, J.** 2004. The function of food-associated calls in white-faced capuchin monkeys (*Cebus capucinus*) from the perspective of the signaller. *Animal Behaviour*, **67**, 431–440.
- Gyger, M. & Marler, P.** 1988. Food-calling in the domestic fowl, *Gallus gallus*: the role of external referents and deception. *Animal Behaviour*, **36**, 358–365.
- Hauser, M. D. & Marler, P.** 1993a. Food-associated calls in rhesus macaques (*Macaca mulatta*). 1. Socioecological factors. *Behavioral Ecology*, **4**, 194–205.
- Hauser, M. D. & Marler, P.** 1993b. Food-associated calls in rhesus macaques (*Macaca mulatta*). 2. Costs and benefits of call production and suppression. *Behavioral Ecology*, **4**, 206–212.
- Hauser, M. D. & Wrangham, R. W.** 1987. Manipulation of food calls in captive chimpanzees: a preliminary report. *Folia Primatologica*, **48**, 207–210.
- Hauser, M. D., Teixidor, P., Field, L. & Flaherty, R.** 1993. Food-elicited calls in chimpanzees: effects of food quantity and divisibility. *Animal Behaviour*, **45**, 817–819.
- Heinrich, B. & Marzluff, J. M.** 1991. Do common ravens yell because they want to attract others? *Behavioral Ecology and Sociobiology*, **28**, 13–21.
- Hohmann, G. & Fruth, B.** 1994. Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology*, **15**, 767–782.
- Janmaat, K. R. L., Byrne, R. W. & Zuberbühler, K.** 2006a. Evidence for a spatial memory of fruiting states of rainforest trees in wild mangabeys. *Animal Behaviour*, **72**, 797–807.
- Janmaat, K. R. L., Byrne, R. W. & Zuberbühler, K.** 2006b. Primates take weather into account when searching for fruits. *Current Biology*, **16**, 1232–1237.
- Lieberman, P.** 1994. Human language and human uniqueness. *Language and Communication*, **14**, 87–95.
- Marler, P.** 1977. Primate vocalization: affective or symbolic? In: *Progress in Ape Research* (Ed. by G. Bourne), pp. 85–96. New York: Academic Press.
- Marler, P. & Tenaza, R.** 1977. Signalling behavior of apes with special reference to vocalisation. In: *How Animals Communicate* (Ed. by T. A. Sebeok), pp. 965–1033. Bloomington: Indiana University Press.
- Marler, P., Dufty, A. & Pickert, R.** 1986. Vocal communication in the domestic chicken. 1. Does a sender communicate information about the quality of a food referent to a receiver? *Animal Behaviour*, **34**, 188–193.
- Marler, P., Evans, C. S. & Hauser, M. D.** 1992. Animal signals: motivational, referential, or both? In: *Nonverbal Vocal Communication: Comparative and Developmental Approaches* (Ed. by H. Papoušek, U. Jürgens & M. Papoušek), pp. 66–86. Cambridge: Cambridge University Press.
- Mundry, R. & Sommer, C.** 2007. Discriminant function analysis with non-independent data: consequences and an alternative. *Animal Behaviour*, **74**, 965–976.
- Newman, J. A. & Caraco, T.** 1989. Co-operative and non-co-operative bases of food-calling. *Journal of Theoretical Biology*, **141**, 197–209.
- Pulliam, H. R. & Caraco, T.** 1984. Living in groups: is there an optimal group size? In: *Behavioural Ecology: an Evolutionary approach* (Ed. by J. R. Krebs & N. B. Davies), pp. 122–147. Sunderland, Massachusetts: Sinauer.
- Roush, R. S. & Snowdon, C. T.** 2000. Quality, quantity, distribution and audience effects on food calling in cotton-top tamarins. *Ethology*, **106**, 673–690.
- Ruxton, G. D.** 1995. Foraging on patches: are groups disadvantaged? *Oikos*, **72**, 148–150.
- Slocombe, K. E. & Zuberbühler, K.** 2005. Functionally referential communication in a chimpanzee. *Current Biology*, **15**, 1779–1784.
- Slocombe, K. E. & Zuberbühler, K.** 2006. Food-associated calls in chimpanzees: responses to food types or food preferences? *Animal Behaviour*, **72**, 989–999.
- Slocombe, K. E. & Zuberbühler, K.** 2007. Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences, U.S.A.*, **104**, 17228–17233.
- Stokes, A. W. & Williams, H. W.** 1971. Courtship feeding in gallinaceous birds. *Auk*, **88**, 543–559.
- Tabachnick, B. G. & Fidell, L. S.** 2001. *Using Multivariate Statistics*, 4th edn. Boston: Allyn & Bacon.
- Van Krunkelsven, E., Dupain, J., Van Elsacker, L. & Verheyen, R. F.** 1996. Food calling by captive bonobos (*Pan paniscus*): an experiment. *International Journal of Primatology*, **17**, 207–217.
- de Waal, F. B. M.** 1988. The communicative repertoire of captive bonobos compared to that of chimpanzees. *Behaviour*, **106**, 183–251.
- Zuberbühler, K.** 2008. Audience effects. *Current Biology*, **18**, R189–R190.