Published in The Journal of the Acoustical Society of America, Vol. 114, Issue 2, 2003, p. 1132-1142
which should be used for any reference to this work

The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization[a)]

Tobias Riede[b)]
Institut für Theoretische Biologie, Humboldt-Universität zu Berlin, Invalidenstrasse 43, 10115 Berlin, Germany

Klaus Zuberbühler[c)]
School of Psychology, University of St. Andrews, St. Andrews, Fife KY16 9JU, Scotland, United Kingdom

Mammalian vocal production mechanisms are still poorly understood despite their significance for theories of human speech evolution. Particularly, it is still unclear to what degree mammals are capable of actively controlling vocal-tract filtering, a defining feature of human speech production. To address this issue, a detailed acoustic analysis on the alarm vocalization of free-ranging Diana monkeys was conducted. These vocalizations are especially interesting because they convey semantic information about two of the monkeys' natural predators, the leopard and the crowned eagle. Here, vocal tract and sound source parameter in Diana monkey alarm vocalizations are described. It is found that a vocalization-initial formant downward transition distinguishes most reliably between eagle and leopard alarm vocalization. This finding is discussed as an indication of articulation and alternatively as the result of a strong nasalization effect. It is suggested that the formant modulation is the result of active vocal filtering used by the monkeys to encode semantic information, an ability previously thought to be restricted to human speech.

I. INTRODUCTION

Male Diana monkeys produce acoustically distinct alarm vocalization to two of their natural predators, the crowned eagles and the leopards (Zuberbühler *et al.*, 1999; Zuberbühler, 2000a). Field playback experiments have shown that nearby listeners respond to these alarm vocalization by producing their own corresponding alarm vocalizations and by showing characteristic locomotor responses. For instance, when hearing leopard alarm vocalization from a concealed speaker, nearby monkeys tend to approach the site of the suspected predator while continuously producing alarm vocalizations, presumably to signal detection to the predator and futility of further hunting (Zuberbühler, Jenny, and Bshary, 1999). Playback experiments have shown that the monkeys' response is driven by the associated meaning, rather than the vocalizations' mere acoustic features (Zuberbühler *et al.*, 1999). This has been taken to suggest that Diana monkey alarm vocalization are another example of natural semantic communication in animals (Seyfarth and Cheney, 2003).

Comparably little is known about the mechanisms of sound production that underlie these behavioral patterns. There is an increasing consensus among researchers in the field that the source-filter-theory, originally put forward to explain speech production (Fant, 1960), serves as a useful model for mammalian sound production (Andrew, 1976; Owren and Bernacki, 1988; Riede and Fitch, 1999). The theory posits that a vocal signal is produced by the vocal folds (the source), and is subsequently shaped by the resonance properties of the vocal tract (the filter). A number of recent studies suggested that some nonhuman species are capable of vocal-tract filtering by controlling the resonance properties independently of the glottal source (Hauser *et al.*, 1993; Hauser and Schön-Ybarra, 1994; Fitch and Reby, 2001). Owren (1990a,b) showed that eagle, and snake alarm vocalizations of vervet monkeys could be distinguished by measures associated with the source, the filter, and timing. They used synthetic versions of these vocalizations to show that individual subjects based their discrimination on acoustic cues associated with the filter, independent of those associated with the source or timing.

The sound production systems of all mammals exhibit a number of fundamental anatomical and acoustical similarities. The primary acoustic signal is generated at a source, typically the vocal folds of the larynx (the glottal source), which are driven into rapid mechanical oscillations by an expiratory airflow from the lungs. The oscillating vocal folds modulate the airflow through the glottal opening, i.e., the airspace between the vocal folds, producing a time-varying acoustic signal: the glottal source signal. The vocal folds are set into vibrations by the combined effect of subglottal pressure, the viscoelastic properties of the folds, and the Bernoulli effect. The aerodynamic energy is sustained by the subglottal pressure, which is maintained by the muscles of expiration. Recently, it has been shown that the vocal folds constitute a highly nonlinear self-oscillating system best modeled as coupled oscillators (Herzel *et al.*, 1995), resulting in the occurrence of nonlinear phenomena in the vocal repertoire.

[a)] Dedicated to Prof. Günter Tembrock on the occassion of his 85th birthday.

[b)] Present address: Medical Science Program, 315 Jordan Hall, Indiana University, IN 47405. Electronic mail: tobiasriede@web.de

[c)] Electronic mail: kz3@st-and.ac.uk

Nonlinear phenomena have been demonstrated earlier in a number of non-human mammals (Wilden *et al.*, 1998), but they have not been found in male Diana monkeys alarm vocalization (Riede and Zuberbühler, 2003).

All mammals have a supralaryngeal vocal tract (hereafter referred to as ‘‘vocal tract’’) through which the sound generated at the glottal source must pass. Like any tube of air, the air column contained in the vocal tract has resonant modes, which selectively allow certain frequencies in the glottal source to pass and radiate through the mouth or nostrils into the environment. The vocal tract hence acts as a bank of bandpass filters, each of which allows a narrow range of frequencies to pass. The vocal-tract resonances, along with the spectral peaks they produce in the vocal signal, have been termed ‘‘formants.’’ Originally, the term was used to describe speech signals (e.g., Fant, 1960; Titze, 1994), but various researchers have used it to describe animal sounds (Lieberman *et al.*, 1969; Nowicki, 1987; Owren and Bernacki, 1988, 1998; Fitch, 1997; Riede and Fitch, 1999). However, many studies have used the term simply to describe spectral concentrations of acoustic energy that appear to be harmonically unrelated to the fundamental frequency (e.g., Hauser *et al.*, 1993; Zuberbühler, 2000a). Unequivocal identification of formants requires an analysis technique that separates the effect of the glottal source from the effect of the vocal tract (e.g., Owren and Bernacki, 1998).

In this study, we investigate the impact of source and vocal-tract parameters on the acoustic structure of Diana alarm vocalizations. We were particularly interested in the question of whether the source-filter theory is suitable to explain the acoustic patterns produced by the Diana monkeys in response to the two predators. We conducted an acoustic analysis based on a linear predictive coding algorithm to determine (i) whether source and tract parameters are independent of each other and (ii) which of the two the monkeys use to convey predator information.

II. METHODS

A. Study site, subjects, and vocalization sample

Data were collected in the Taï National Park, Côte d’Ivoire, between June 1994 and June 1997 in an approximately 40-km^2 study area of primary rain forest surrounding the Centre de Recherche en Ecologie (Université de Abobo-Adjame, Abidjan) research station (5° 50′ N, 7° 21′ W), in the Taï National Park, Côte d’Ivoire. Seven monkey species were regularly observed in the area: the western red colobus (*Colobus badius*), the western black-and-white colobus (*Colobus polykomos*), the olive colobus (*Colobus verus*), the Diana monkey (*Cercopithecus diana*), the lesser white-nosed monkey (*Cercopithecus petaurista*), the Campbell’s monkey (*Cercopithecus campbelli*), and the sooty mangabey (*Cercocebus atys*). Diana monkey groups typically consisted of about 20–25 individuals with one adult male and several adult females with their offspring, each occupying and defending a stable home range of less than 1 square kilometer against neighboring groups.

We analyzed the vocalizations of ten different adult males from ten different Diana monkey groups, five responding to an eagle and five to a leopard. We analyzed the first five vocalizations given by each individual, resulting in a sample of 25 eagle and 25 leopard alarm vocalizations. Recordings were made with a Sony WMD6C tape recorder and a Sennheiser microphone (ME88 head with K3U power module) on 90-min type IV metal tapes. The frequency response of the microphone (40 Hz–20 kHz; ±2.5 dB) and the tape recorder (40 Hz–14 kHz, ±3 dB; distortion of 0.1%; signal-to-noise-ratio of 57 dB) are flat and within the frequency range of analysis. Recordings were made at distances ranging from 20 to 50 meters.

B. Acoustic analysis

We digitized all recordings at a 16-bit quantization and a 44-kHz sampling rate using SIGNALIZE software. We performed signal analysis on a PC using the signal-processing software HYPERSIGNAL-MACRO™ and a DSP32C PC system board. We completed the spectrographic analysis by using a 512-point fast Fourier transformation, with 75% frame overlapping, a 44-kHz sampling frequency, and a Hanning window. To avoid aliasing effects we low-passed filtered all vocalization at 22 kHz.

Linear predictive coding (‘‘LPC’’) is a spectral modeling technique used to estimate formant frequencies in human speech. LPC uses least-squares curve fitting to estimate the value of a point in a time-domain waveform based on the past N points, where N is the order of the LPC analysis. LPC algorithms then construct the best-fitting all-pole model to account for the waveform. ‘‘All-pole’’ means that only vocal-tract resonances (‘‘poles’’) are estimated, and not antiresonances (‘‘zeros’’). Such a spectral model appears to be a valid first approximation for most human speech signals and the monkey vocalization analyzed here (Markel and Gray, 1976). The specifics of the algorithms used in linear predictive coding are described elsewhere (see Markel and Gray, 1976 for the mathematical details, and Owren and Bernacki, 1998 for application in bioacoustics). In the current analysis we used the autocorrelation technique provided by HYPERSIGNAL™. The technique outputs the coefficients of an Nth-order all-pole digital filter whose frequency response best approximates in a least-squares sense the spectrum of the input signal. Given a broadband source signal and an appropriate model order [typically estimated as $2+2*N(\text{formants})$], LPC analysis can provide an extremely accurate estimate of formant center frequencies in both human speech and animal sounds. Signal analysis was conducted with HYPERSIGNAL-MACRO™ with 12 coefficients and pre-emphasis settings of 0.8 to 0.99. All LPC measurements were visually verified by superimposing the LPC-derived frequency response over a 512-point fast Fourier transform (FFT) of the same time slice, allowing the user to select the optimum number of coefficients for each vocalization by trial and error.

In order to make formant extraction more reliable, acoustic parameters were interpreted using anatomical estimates. Since the recorded signal is a combination of the primary signal spectrum and the transfer function of the vocal tract, peaks in the spectrum can be the result of the source

TABLE I. Pulse duration. Minimum (P-dur min), maximum (p-dur max), mean (p-dur mean) of pulse duration in eagle (individuals 1 to 5) and leopard (individuals 6 to 10) alarm vocalizations. For each individual 10 calls (*N*) (2 calls per bout, 5 bouts per individual) were measured. In individuals 6 and 7, 1 out of 5 bouts contained only 1 call; therefore, 1 other bout was chosen to deliver 3 calls. In individual 9 only 4 bouts were available, and 1 of those delivered only 1 call.

Individual	*N*	p-durat min (ms)	p-durat max (ms)	p-durat mean (ms)	s.d. (ms)	Jitter (%)
			Eagle alarm			
1	10	13.9	19.9	16.4	1.1	6.4
2	10	8.3	21.9	15.9	0.6	8.7
3	10	13.3	22.2	17.4	1.5	6.5
4	10	8.3	20.3	14.1	1.7	9.2
5	10	10.1	24.0	16.9	1.1	8.3
mean	50			16.1	1.5	7.8
			Leopard alarm			
6	10	13.3	29.9	16.7	2.3	8.9
7	10	14.6	25.6	18.3	2.1	6.9
8	7	12.1	22.7	15.6	2.5	6.8
9	10	15.1	26.7	17.4	1.1	8.1
10	10	14.8	28.3	17.4	1.4	7.4
Mean	47	13.3	29.9	17.1	2.1	7.6

(e.g., regular/harmonic or irregular patterns) as well the vocal tract (i.e., formants). That makes it important to sort out if spectral peaks can be considered as formants. The relationship between vocal-tract length and formant position follows Eq. (1):

$$Fn=[(2n-1)*c]/4\mathrm{VTL}. \qquad (1)$$

In this equation, Fn is the nth formant frequency in Hz; c is speed of sound (35 000 cm/s in warm humidified air) and VTL is vocal tract length in cm. This relationship is true for a uniform, hard-walled tube closed at one end, but it has also shown to be a good approximation for several nonhuman mammalian species' vocal tracts; e.g., domestic cats (Carterette *et al.*, 1979), rhesus macaques (Fitch, 1997; Rendall *et al.*, 1998), and domestic dogs (Riede and Fitch, 1999). Broadband utterances are particularly well suited for formant extractions (Owren and Linker, 1995) because they reflect the vocal-tract transfer function best. In addition to the male Diana monkey vocalization, our focus of interest in this paper, we also examined broadband female vocalizations for formant characteristics. Females are approximately 20% smaller than males, suggesting also to have an at least 20% shorter VTL (Fitch, 1997; Riede and Fitch, 1999).

C. Acoustic parameters

The basic acoustic unit in the Diana monkey alarm vocalization is the pulse (Riede and Zuberbühler, 2003). The pulse, which resembles a damped oscillation, is a rapid-amplitude transition of the signal from a baseline value to a higher or lower value, followed by rapid return to baseline. We measured pulse duration (Table I) as the interval between two pulse onsets. Pulse duration corresponds to the fundamental frequency of the oscillating vocal folds. The next higher acoustic units of Diana monkey alarm vocalizations are calls (Fig. 1), whose duration we also measured. Note that this terminology differs from the one used in a previous study (Zuberbühler, 2000a). A call consists of a series of pulses of varying duration. This cycle-to-cycle variability in fundamental frequency was termed jitter (Table I, Titze, 1994), calculated as the ratio of standard deviation to mean of the pulse duration per call. For each individual a mean jitter was calculated based on ten calls (except in one individual, when only seven calls were available). The final temporal parameter was bout duration, calculated as the overall duration of continuous acoustic energy. As spectral parameters we measured formants and formant bandwidth. Formant bandwidth is the size of the formant in the spectral representation. Two points are identified on the slopes of the resonance curve, where the response is 3 dB lower than at the peak. The difference in frequency between the 3-dB points defines the formant bandwidth. The size of the formant bandwidth is determined by the amount of attenuation in the vocal tract. Statistical analyses were based on monovariate Mann-Whitney *U*-tests, nested analysis of variance and factor analysis (SPSS 10.0).

III. RESULTS

A. Source acoustics

Pulse duration varied in eagle alarm vocalization between 8.3 and 24 ms (mean±s.d. 16.1±1.5), suggesting a fundamental frequency between 42 and 120 Hz (mean 62 Hz). It varied in leopard alarm vocalizations between 13.3 and 29.9 ms (mean±s.d. 17.1±2.1), suggesting a fundamental frequency between 33 and 75 Hz (mean 58 Hz). Pulse duration differed significantly between eagle and leopard alarm vocalizations ($N_1=50$, $N_2=42$, $U=649$, $P<0.001$, Mann-Whitney *U*-test, two-tailed), although the individual values overlapped broadly (Table 1). In both vocalization types, the pulses occurred very regularly, suggesting a rigid vibration pattern of the oscillating glottal source. Jitter did not differ significantly between eagle and leopard alarm vocalization (eagle alarm vocalization: mean 7.8%, leopard alarm vocalization: mean 7.6%; $N_1=5$, $N_2=5$, $U=13$, $P>0.2$, Mann-Whitney *U*-test, two-tailed). Call duration did

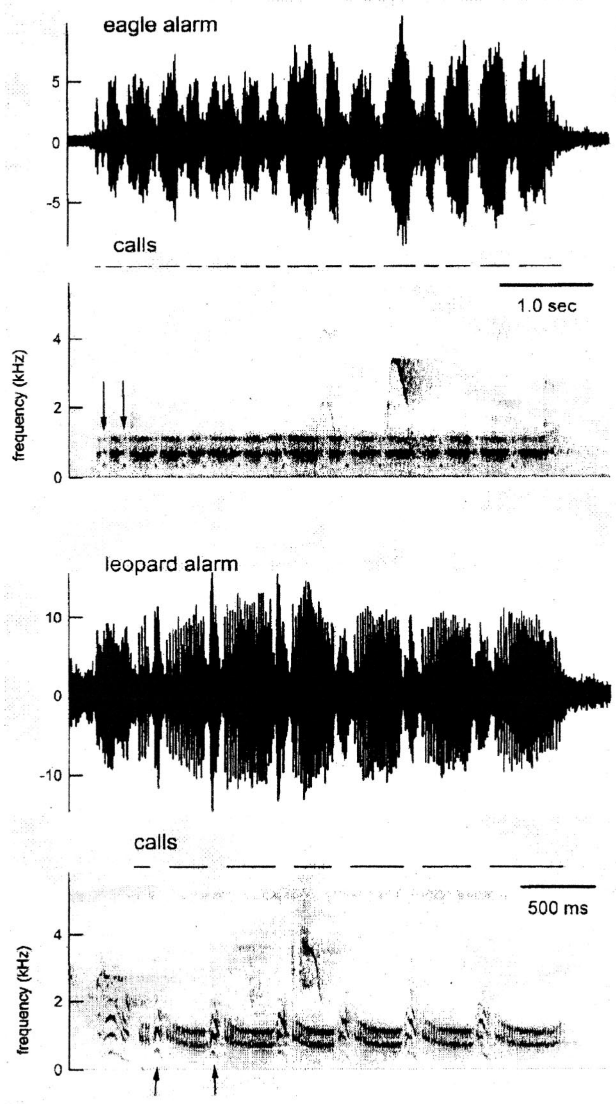

FIG. 1. Example of a male eagle alarm vocalization and a male leopard alarm vocalization. Top: time series, bottom: spectrogram. Calls are indicated by horizontial lines. The first two intercall elements in both bouts are indicated by arrows. In the range between 2 and 5 kHz several structures are visible (arch-like frequency bands), caused by alert calls given by other Diana monkey group members.

not differ significantly between eagle and leopard alarm vocalizations (leopard alarm vocalization: range: 56 and 566 ms; eagle alarm vocalizations range: 23 and 717 ms; $N_1 = 25$, $N_2 = 25$, $U = 287$, $P > 0.3$, Mann-Whitney U-test, two-tailed, Table II). Bout duration differed significantly between eagle (mean±s.d., 3615±1392 ms) and leopard alarm vocalizations (mean±s.d., 2206±1321 ms) ($N_1 = 25$, $N_2 = 25$, $U = 146$, $P < 0.001$), although the individual values overlapped broadly (Table II).

B. Tract acoustics

1. Identifying formants

The average power spectrum of five subsequent pulses of a male leopard alarm vocalizations is depicted in Fig. 2: Two prominent peaks are visible between 600 and 1500 Hz, both identified by LPC analysis. The first peak ranged between 690 and 1000 Hz (Table III). Assuming that this peak has formant properties, then Eq. (1) suggests a vocal-tract length between 8.8 and 12.7 cm, which is anatomically reasonable. There was a significant difference between the $F1$ values at the beginning of the calls in eagle and leopard alarm vocalizations ($N_1 = 50$, $N_2 = 51$, $U = 40$, $P < 0.001$, Mann-Whitney U-test, two-tailed, Table III). LPC analysis depicted a second peak at around 1300 Hz. To consider this peak as a second formant is difficult to reconcile with Eq. (1) since it presupposes a substantially longer vocal tract of approximately 20 cm. The second peak was close to the first one, even during modulation, and it exhibited a broader bandwidth. In some high-quality recordings with low background sound-pressure level, a third peak near 2800 Hz was visible. Assuming that this peak has formant properties (i.e., second formant), then Eq. (1) suggests a vocal-tract length similar to the one predicted by the first peak. In the female alarm vocalization (Fig. 3) two separate peaks were visible at 1000 and at 2800 Hz, both again suggesting a VTL of roughly 9 cm. An additional peak around 1300 Hz was absent in female vocalization.

2. Formant behavior

Leopard and eagle alarm vocalizations differed most strongly in the downward modulation of the first formant ($\Delta F1$-start; Figs. 1 and 2 and Table IV). Although present in both alarm vocalization types, the modulation was three times stronger in the leopard alarm vocalization and there was little overlap between vocalization types (Table IV; $N_1 = 50$, $N_2 = 51$, $U = 174$, $P < 0.001$, Mann-Whitney U-test, two-tailed), and although the variability between individuals is significant, the nested ANOVA indicates that the variability between "eagle" and "leopard" calls in the parameter "formant transition" is even higher (Table V).

Simultaneously, the second prominent peak at around 1300 Hz modulated downwards, but to a lesser extent. There was a significant difference between the first and the second peak difference at the beginning and in the middle of the call ($T = 1064$; $P < 0.05$, Wilcoxon test, two-tailed), suggesting that the first peak modulated stronger than the second one. Although $\Delta F1$-start was defined as the difference between the beginning and the middle of the call, the actual modulation reliably occurred during the first four to six pulses, corresponding to less than 20% of all pulses in the entire call (Fig. 4). Finally, the modulation of $F1$ between the middle and the end of the call was not significantly different between two alarm vocalization types, and individual values overlapped strongly ($N_1 = 50$, $N_2 = 50$, $U = 1166$, $P > 0.05$, Mann-Whitney U-test, two-tailed). A factor analysis indicated that 35% of overall acoustic variability was explained by the formant downward modulation.

IV. DISCUSSION

In this study we were interested in the acoustic structure of Diana monkey alarm vocalization and in how the various acoustic parameters segregated eagle from leopard alarm vocalizations. Our analyses suggested that the spectral peaks in the alarm vocalizations have formant properties, i.e., that the spectral concentrations of acoustic energy are harmonically

TABLE II. Acoustic parameters of eagle and leopard alarm vocalizations. Minimum, maximum, and mean number of calls (no call min; no call max; no call mean); minimum, maximum, and mean of the call duration (call dura min, call dura max, call dura mean) and bout duration of 5 individuals each represented with 5 Eagle and 5 Leopard alarm bouts.

	No call min	No call max	No call mean (ms)	Call dura min (ms)	Call dura max (ms)	Call dura mean (ms)	Bout duration (ms)
			Eagle Alarm				
Ind.1	4	7	5.8	64	717	294	2107
Ind.2	8	15	10.6	58	515	256	3593
Ind.3	10	16	13.6	23	495	242	4439
Ind.4	6	19	11.0	39	604	232	3490
Ind.5	10	23	15.0	56	472	199	4446
Mean			11.2			236±122 $N=279$	3615±1392 $N=25$
			Leopard alarm				
Ind.6	1	17	8.6	56	482	264	3308
Ind.7	1	13	5.2	71	424	213	1584
Ind.8	2	6	3.6	73	369	221	1066
Ind.9	3	12	6.4	73	506	197	2038
Ind.10	6	9	7.2	73	566	293	3030
Mean			6.2			243±116 $N=155$	2205±1321 $N=25$

unrelated to the fundamental frequency. A formant downward modulation at the beginning of the call most reliably distinguished between the two alarm vocalization types because there was little to no overlap between individual vocalizations, monovariate statistical analyses yielded the highest *P* values, and a factor analysis showed that 35% of overall acoustic variability could be explained by this one parameter.

A. Source acoustics

Although source acoustic parameters and temporal cues (pulse duration, call duration, bout duration, number of calls per bout, jitter) differentiated eagle and leopard alarm vocalization to various degrees, they were unable to distinguish between the two vocalization types: the parameter distribution was strongly overlapping between both types of vocalization.

Our analyses showed that the pulse character of the alarm vocalization was surprisingly consistent. Earlier studies considering the vocal source as a nonlinear system found much higher degrees of irregularities in vocal utterances. For example, up to 15% of human infant cries and animal vocalization contained nonlinear phenomena (Robb and Saxman, 1988; Wilden *et al.*, 1998; Riede *et al.*, 2000). Nonlinear phenomena were virtually absent in male Diana monkey vocalizationa and pulses were not interrupted by any other vibration modes of the vocal folds, except for the short harmonic intercall elements. Specialized system adjustments and the low fundamental frequency could account for the remarkable stability of the oscillating system (but see Mergell *et al.*, 1999).

The spectrum of a pulse at the source is expected to be broadband (Titze, 1994; Au, 1993). This suggests that prominent peaks in the spectrum of the emitted signal are due to filtering effects in the vocal tract. Because of their broad bandwidth, pulses are well suited to picture the resonance characteristics of the vocal tract. This characteristic, as well as the robustness of this source signal and the very low fundamental frequency let Riede and Zuberbühler (2003) to hypothesize that male Diana monkey alarm calls are adapted

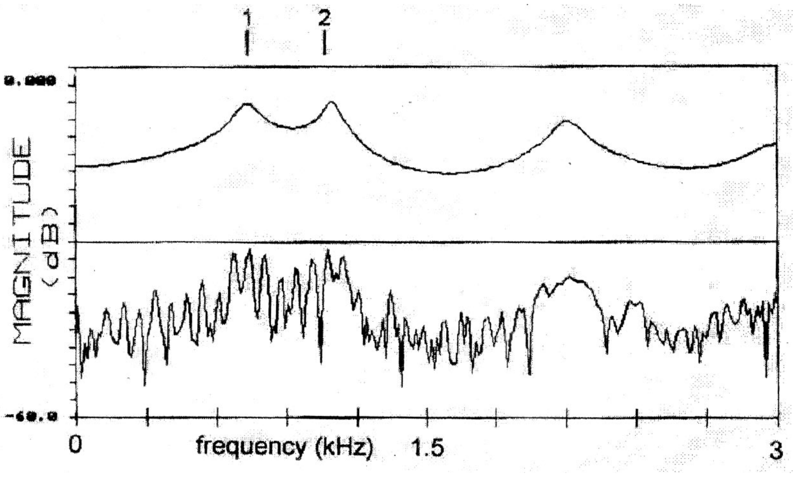

FIG. 2. Spectrum (lower curve) and LPC spectrum (upper curve) of a 100-ms segment of a male eagle alarm call. Two peaks are indicated (“1” and “2”). A third peak around 2100 Hz does not belong to the monkey call.

TABLE III. First formant at the beginning of the call. Mean and standard deviation of the first peak ($F1$) and the second peak ($F2$) at the beginning of the call ($F1$-start±s.d. and $F2$-start±s.d.). Mean and standard deviation of the bandwidth of the first ($B1\pm$s.d.) and second ($B2\pm$s.d.) peak at the beginning of the call.

	N	$F1$-start±s.d. (Hz)	$F2$-start±s.d. (Hz)	$B1\pm$s.d. (Hz)	$B2\pm$s.d. (Hz)
		Eagle alarm			
Individual 1	10	753±52	1298±45	132±53	135±52
Individual 2	10	744±18	1194±20	71±17	117±36
Individual 3	10	692±19	1221±238	69±19	109±41
Individual 4	10	824±26	1227±134	95±25	102±24
Individual 5	10	802±47	1287±82	88±29	120±34
Mean		763±58	1246±130	91±38	117±38
		Leopard alarm			
Individual 6	10	916±67	1371±155	69±16	154±55
Individual 7	10	947±46	1339±68	87±15	90±37
Individual 8	7	907±36	1466±68	86±40	181±35
Individual 9	10	996±71	1473±123	78±32	159±39
Individual 10	10	944±53	1447±142	80±25	167±40
Mean		944±63	1407±125	80±26	150±51

for a more elaborated vocal tract performance, i.e. formant variability. Preliminary video studies suggest a fourth characteristics of this system. There is a distinct possibility that the alarm calls investigated in this study are product of vocal fold vibration caused by air flowing *into* the lungs, rather than the more commonly observed expiratory sound production mechanism in mammals. In this case, the brief but highly harmonic inter-call elements seen in the spectrograms (Fig. 1) would have to be considered the product of vocal fold vibration caused by expiration. Inspiration-caused sound production could also explain the remarkable and otherwise rarely observed absence of nonlinear phenomena during call-

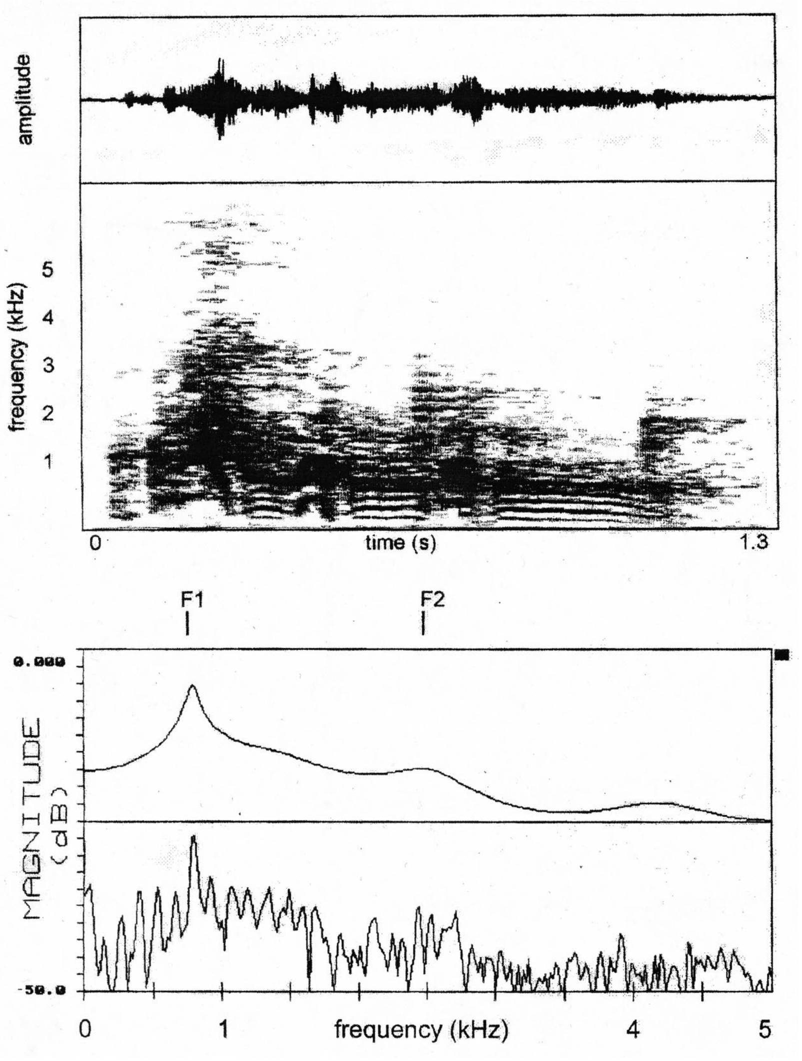

FIG. 3. Example of a female eagle alarm vocalization. Top: time series, middle: spectrogram. Bottom: a 100-ms segment is given in the frequency domain (lower curve) plus a LPC spectrum is given (upper curve). Two peaks are indicated ($F1$,$F2$) representing the first and second formant.

TABLE IV. Formant characteristic of eagle and leopard alarm calls. Mean ($\Delta F1$-start)±standard deviation (s.d.) of the down modulation of the first formant measured between the beginning and the middle of the call. Mean ($\Delta F1$-end)±standard deviation (s.d.) of the modulation of the first formant measured between the middle and the end of the call. For each individual a certain number of calls (N) were considered for measurements.

	N	$\Delta F1$-start±s.d. (Hz)	$\Delta F1$-end±s.d. (Hz)
		Eagle alarm	
Individual 1	10	55.8±60	8.5±33
Individual 2	10	21.6±14	−2.1±23
Individual 3	10	23.5±25	−4.3±22
Individual 4	10	66.5±23	−4.3±28
Individual 5	10	36.5±30	−10.7±18
Mean		50.8±34	−2.5±25
		Leopard alarm	
Individual 6	10	187.3±45	25.9±95
Individual 7	10	187.3±59	19.2±56
Individual 8	7	89.9±46	21.5±32
Individual 9	10	128.0±65	−16.4±23
Individual 10	10	185.0±72	6.6±46
Mean		153.5±68	11±55

ing, perhaps the product of a highly stable oscillating system. High resolution video analyses of the respiratory movements of the thorax during calling will be necessary to resolve this issue. Although we expect that the actual direction of airflow during call production has no direct implications for the explanation of the formant nature of the calls (Robb *et al.* 2001), the evolutionary insinuations would be quite intriguing. What selection pressures could have lead to the invention of an inspiration-based semantic communication system in the evolutionary history of this taxon?

B. Tract acoustics

Equation (1) makes a number of specific predictions regarding the location of the different formant frequencies. A crucial element in the equation is vocal-tract length, a parameter for which no anatomical data are available at the moment. However, previous studies have shown that in mammals there is strong positive correlation between skull length and vocal-tract length, and (in resting position of the larynx) that VTL is maximally as long as skull length but mean VTL is shorter than mean skull length (Fitch, 1997; Riede and Fitch, 1999; Fitch, 2000). Table VI lists skull lengths for Diana monkeys from various museum specimens, measured as the distance between the front of the incisor teeth and the *Protuberantia occipitalis*. Based on the relationship found by Fitch (1997) and Riede and Fitch (1999) the skull data (Table VI) suggest a vocal-tract length of maximally 11.5 cm, which predicts [according to Eq. (1)] a first formant at around 760 Hz, and a second formant at around 2280 Hz. LPC analysis depicts a peak in male and female vocalization near 800 Hz, strongly suggesting that this peak is the equivalent of a first formant. The 11.5-cm VTL also predicts [according to Eq. (1)] a second formant around 2280 Hz. LPC depicts a second peak at 2800 Hz, in males only in high-quality recordings (close distance between microphone and vocalizer, and very low background noise), and more regular in female vocalizations. This suggests that this peak is possibly the equivalent of the second formant. LPC depicts regularly a peak around 1300 Hz in male alarm vocalization, and it will be discussed below how this peak could be explained.

TABLE V. Nested ANOVA of the relationship between alarm call type (eagle versus leopard) and individual specificity.

	DF	Sum of squares	Mean square	F-ratio	Probability level
Call type	1	342 381	342 381	33.45	<0.001
Individual identity	8	81 880	10 235	4.41	<0.001
S	87	202 142	2 323		
Total	96	626 404			

C. The effects of changes in vocal tract diameter

Our calculations are based on the assumption that the formants are created in a uniform tube with no significant changes in tube diameter. It is well known that deviation from such uniformity often results in formant shifts. In humans, nonuniformity of the vocal tract is common, and speech vowels are prominent acoustic products of such nonuniformity (Fant, 1960). The consistent peaks of male Diana monkey vocalization at around 1300 Hz (see Fig. 3), therefore, could be the result of changing tube diameters in the male vocal tract. It is interesting that to the human ear, the calls of an eagle alarm vocalization strongly resemble the human vowel /o/. In contrast, the calls of a leopard alarm vocalization strongly resemble a vowel transition from /ʌ/ to /o/. In human speech, the second formants of both /ʌ/ and /o/ are lowered and they approach the first formant (Story *et al.*, 1996), similar to the monkey vocalizations. Therefore, one interpretation of these data is that, similar to human speech sounds, the second peak in male alarm vocalization around 1300 Hz represents the second formant ($F2$), due to variation in vocal-tract diameter and hence deviating from the predictions made by Eq. (1). Detailed anatomical work will be necessary to resolve this issue.

D. The effects of nasalization

An alternative explanation for the presence of acoustic energy at around 1300 Hz in male alarm vocalizations is provided by nasalization. Nasalization is produced by coupling the oral and nasal cavities via the velopharyngeal opening (*Ostium intrapharyngeum*). During this process the velum is lowered, resulting in coupling between the nasal cavity and the oral vocal tract. Acoustically, nasalization replaces the sharp spectral peak of the first formant ($F1$) by two peaks, the oral and the nasal pole (Fant, 1960), which widens the bandwidth of the first formant (Dang and Honda, 1996). Hence, nasalization shifts the natural frequencies of the oral part of the vocal tract by adding pole-zero pairs to the vocal-tract transfer function. These acoustic effects are strongest at low frequencies, in the vicinity of the first formant (House and Stevens, 1956; Stevens, Fant, and Hawkins, 1987; Maeda, 1993; Dang and Honda, 1996). In the higher frequencies, nasalization may introduce shifts in formants, modification in formant amplitudes, and additional spectral peaks. However, these effects are not as consistent

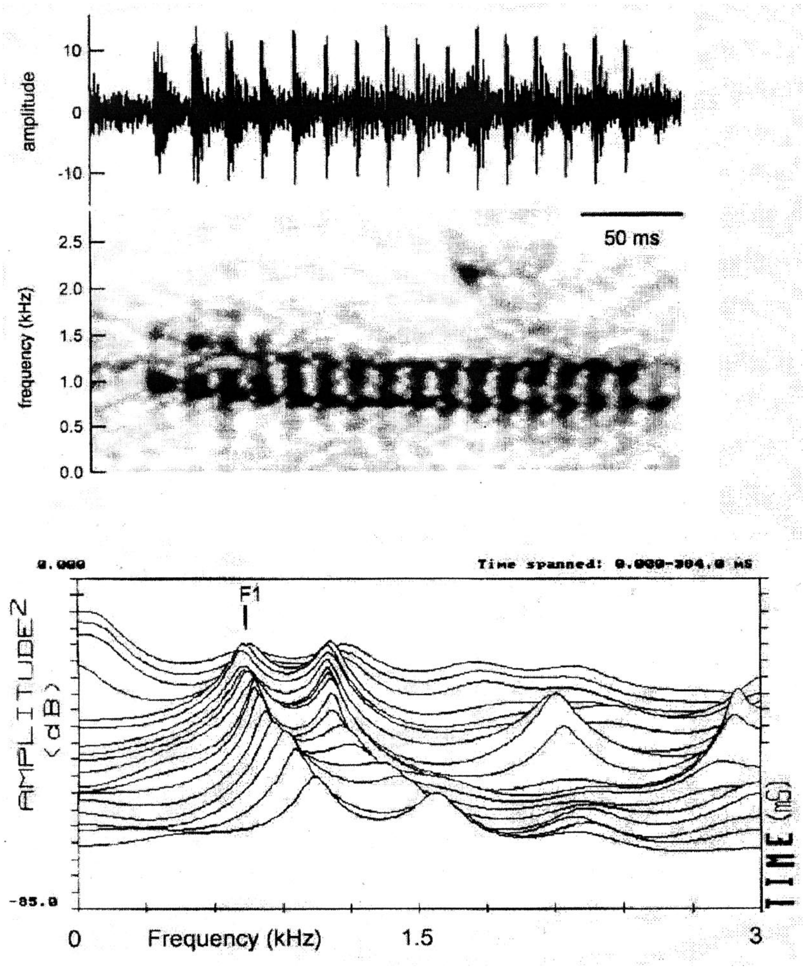

FIG. 4. Example of a male leopard alarm call. From top to bottom: time series, spectrogram, and waterfall representation of the LPC curves. The bottom waterfall representation is a three-dimensional display and shows the Fourier transform spectra of several time slices, altogether 0.384 s. The first formant ($F1$) is indicated in the waterfall representation.

as those in the vicinity of the first formant. Acoustic modeling showed that nasalization can also have secondary spectral effects, for example by reducing the amplitude of higher frequencies, i.e., second and third formants, possibly due to the strong dampening in the nasal tract (House and Stevens, 1956, p. 225).

In acoustic studies on nonhuman mammals, nasalization has been used to explain the appearance of subharmonics in the spectrogram of rhesus monkey (*Macaca mulatta*) vocalizations (Hauser, 1992). However, this interpretation is controversial because other work has shown that subharmonics are considered nonlinear phenomena, caused by vocal-fold vibration (Wilden *et al.*, 1998). Similarly, selective dampening of the fundamental frequency of siamang vocalization has been attributed to nasalization (Haimoff, 1983). However, an alternative explanation for this finding is that the zeros of the oral vocal tract were responsible for filtering the source signal in these animals. In Diana monkey alarm vocalizations, however, the two spectral peaks visible in the range of the first formant might be the result of nasalization. The lower peak around 800 Hz likely represents the first oral tract resonance, while the second peak around 1300 Hz could be the result of the first nasal resonance ($F1_n$). Thus, in Diana monkeys, nasalization can readily explain (i) the lack of acoustic energy above 1500 Hz due to dampening by the nasal tract; (ii) the close proximity of the first two peaks due to its tendency to replace the first formant by two peaks, the oral and the nasal pole.

TABLE VI. Skull length measurments (in cm) from skulls

	Skull length (cm)
1 Adult male (no: 2578)[a]	12.1
4 Adult males[b]	11.2±0.9
1 Adult female (no: 2620)[a]	10.7
8 Adult females[b]	9.9±0.3
2 Adult females[c]	both 10.3
4 Adult males[d]	11.5±0.54
3 Adult females[d]	9.84±1.1

[a] From the Paris Natural History museum, Laboratoire Mammifères et Oiseaux.

[b] http://1kai.dokkyomed.ac.jp/mammal/en/species/cercopithecus diana.html

[c] From the Indiana University at Bloomington Dept. of Anthropology collection.

[d] From the American Museum of Natural History at New York.

E. The effect of extralaryngeal cavities

It has been suggested that additional oral cavities branching off the oral vocal tract could affect formant behavior. For example, various guenons are known to possess pairwise or singular forms of air sacs, which develop directly from the laryngeal or pharyngeal cavity (Gautier, 1971). Perforation of air sacs reduced the amplitude of the vocalization, enriched the spectral pattern because more harmonics were visible in the spectrogram, and introduced irregular noise (Gautier, 1971). Recent modeling work on the effects of adding a side branch (like the air sac) to the oral vocal tract leads to the consistent introduction of zeros into the transfer function and to an overall spectrum that resembled that of the Diana monkey alarm vocalizations (Jackson *et al.*, 2001).

F. Formant frequency modulations

Leopard and eagle alarm vocalizations differed most prominently in the downward modulation of the first formant at the beginning of each call (Fig. 1 and Table IV) with leopard alarm vocalizations exhibiting a threefold stronger downward modulation than eagle alarm vocalizations. The strength of downward modulation was very consistent and differentiated the two alarm vocalizations exclusively, suggesting that one single parameter differentiated between the two alarm vocalization types to a large degree. The formant differences at the beginning of a call indicated different configurations of the vocal tract, for instance in relative length, cross-sectional areas, or coupling with the nasal tract.

To human listeners the downward modulation of the first formant provide the perceptual impression of a transition from a /ʌ/ vowel to an /o/-like vowel. Anatomically the /ʌ/–/o/ transition is mainly correlated to an increase in volume of the pharyngeal chamber with minimal vocal-tract elongation (Story *et al.*, 1996). In this transition both formants modulate downwards, although the first formant does it more strongly than the second one. According to Peterson and Barney (1952) and Story *et al.* (1996), the /ʌ/ to /o/ transition is caused by a first and second formant decrease, similar to the findings in male Diana monkey alarm vocalizations. In the monkey alarm vocalizations both peaks modulated downwards simultaneously, while the first one did so more strongly than the second one. The articulatory maneuver in the frontal oral cavity responsible for the shift from /ʌ/ to /o/ appeared to be similar in the leopard alarm vocalizations, namely a lifting of the mandible combined with protrusion of the lips. Mandible lifting by female rhesus macaques (*Macaca mulatta*) caused decreases in the dominant frequencies of coo vocalization (Hauser *et al.*, 1993), although the study did not distinguish between dominant frequencies and formant frequencies.

A number of additional mechanisms could be causally related with the observed formant downward modulation. Vocal-tract elongation has the effect of decreasing formant frequencies [see Eq. (1)], either by lip protrusion or by larynx lowering, as recently demonstrated for red deer (Fitch and Reby, 2001).

G. Perceptual salience of the observed variation in formant modulation

Although no direct evidence is available, it is reasonable to assume that Diana monkeys are sensitive to variation in formant behavior and able to use this acoustic cue as a base for important behavioral decisions for the following reasons. First, humans classify vowels primarily on the basis of the two lowest formant frequencies (Peterson and Barney, 1952; Bogert and Peterson, 1957; Kent, 1978, 1979), suggesting that Diana monkey alarm vocalization provide sufficient acoustic information for accurate discriminations. Second, various birds and mammals have been tested successfully on their discriminative abilities on human vowels (baboons, Hienz and Brady, 1988; dogs, Baru, 1975; cats, Dewson, 1964; blackbirds and pigeons, Hienz *et al.*, 1981). For example, Japanese macaques (*Macaca fuscata*) were tested on single- and multiformant tone complexes, revealing formant-frequency discrimination abilities at 500 Hz and 1.4 kHz (Sommers *et al.*, 1992), which indicates that the formant downward modulation in Diana monkey alarm vocalization is perceptually salient to primate recipients. In chacma baboons (*Papio cyncephalus ursinus*) (Hienz and Brady, 1988) and Japanese macaques (*Macaca fuscata*) (Sinnott, 1989) it was shown that individuals readily discriminated among typical American-English vowel sounds. Japanese macaques and Sykes monkeys (*Cercopithecus albogularis*) were as good as human subjects in this task (Sinnott and Kreiter, 1993; Sinnott *et al.*, 1997). Sinnott and Kreiter (1993) synthesized a steady-state vowel continuum (formant behavior similar in appearance as in the Diana monkey leopard alarm vocalizations) by varying the first and second formant. Subjects heard a given vowel background and responded to changes towards a second vowel target, and monkeys showed similar sensitivity like humans (Sinnott and Kreiter, 1993). Some studies suggested that, similar to humans, nonhuman primates give the first formant more importance than the second formant (Kojima and Kiritani, 1989; Brown and Sinnott, 1990; Sinnott *et al.*, 1997).

Several studies have suggested that formant frequencies are likely to play an important role in nonhuman primate communication. Owren and Bernacki (1988) and Owren (1990c) used operant techniques to demonstrate that spectral features potentially related to formants are discriminated by vervet monkeys in their own species-specific vocalizations. In baboon vocalizations, Owren *et al.* (1997) found a relationship between changes of formant-related spectral peak patterns and social context. Rendall *et al.* (1999), looking at baboon grunt vocalizations, found that formant characteristics were correlated with social context. Rhesus macaque coo vocalizations and grunts are the effects of vocal-tract filtering, but these vocalizations are more likely to be related to individual differences than external events (Rendall *et al.*, 1998). This study reveals that in certain contexts nonhuman primates are able to engage vocal-tract changes to encode important events in the environment, a defining feature of human speech production.

ACKNOWLEDGMENTS

We thank Michael Owren, Robert Seyfarth, and one anonymous reviewer for comments on an earlier version of the manuscript. We are grateful to Gillian Page for providing us with video footage of male calling behavior. This work was supported by a fellowship within the Postdoc Program of the German Academic Exchange Service (DAAD) to TR.

Au, W. (**1993**). *The Sonar of Dolphins* (Springer, New York).

Andrew, R. J. (**1976**). ‘‘Use of formants in the grunts baboons and other nonhuman primates,’’ Ann. (N.Y.) Acad. Sci. **280**, 673–693.

Baru, A. V. (**1975**). ‘‘Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration and number of formants) in dog,’’ in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. A. A. Tatham (Academic, New York).

Bogert, B. P., and Peterson, G. E. (**1957**). ‘‘The acoustics of speech,’’ in *Handbook of Speech Pathology*, edited by L. E. Travis (Appleton-Century-Cofts, New York), pp. 109–173.

Brown, C. H., and Sinnott, J. M. (**1990**). ‘‘The perception of complex acoustic patterns in noise by Blue monkey (*Cercopithecus mitis*) and human listeners,’’ Int. J. Comp. Psychology **4**, 79–90.

Brown, C. H., and Cannito, M. P. (**1995**). ‘‘Modes of vocal variation in Sykes’s monkey (*Cercopithecus albogularis*) squeals,’’ J. Comp. Psychol. **109**, 398–415.

Carterette, E. C., Shipley, C., and Buchwald, J. S. (**1979**). ‘‘Linear prediction theory of vocalization in cat and kitten,’’ in *Frontier of Speech Communication Research*, edited by B. Lindblom and S. Öhman (Academic, New York), pp. 245–257.

Dang, J., and Honda, K. (**1996**). ‘‘A new method for measuring antiresonance details of the vocal transmission characteristics: An experimental study of acoustic tubes,’’ J. Acoust. Soc. Jpn. (E) **17**, 93–99.

Dewson, J. H. (**1964**). ‘‘Speech sound discrimination by cats,’’ Science (Washington, DC, U.S.) **141**, 555–556.

Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Fitch, W. T. (**1997**). ‘‘Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques,’’ J. Acoust. Soc. Am. **102**, 1213–1222.

Fitch, W. T. (**2000**). ‘‘Skull dimensions in relation to body size in nonhuman mammals: The causal bases for acoustic allometry,’’ Zoology **103**, 40–58.

Fitch, W. T., and Reby, D. (**2001**). ‘‘The descended larynx is not uniquely human,’’ Proc. R. Soc. London, Ser. B **268**, 1669–1675.

Gautier, J. P. (**1971**). ‘‘Etude morphologique et fonctionnelle des annexes extra-laryngées des cercopithecinae; liason avec les cris d’espacement,’’ Biol. Gabon. **7**, 229–267.

Haimoff, E. H. (**1983**). ‘‘Occurrences of anti-resonance in the song of the siamang (*Hylobates syndactylus*),’’ Am. J. Primatol **5**, 249–256.

Hauser, M. D. (**1992**). ‘‘Articulatory and social factors influence the acoustic structure of rhesus monkey vocalization: A learned mode of production?’’ J. Acoust. Soc. Am. **91**, 2175–2179.

Hauser, M. D., Evans, C. S., and Marler, P. (**1993**). ‘‘The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalization,’’ Anim. Behav. **45**, 423–433.

Hauser, M. D., and Schön-Ybarra, M. (**1994**). ‘‘The role of lip configuration in monkey vocalization: Experiments using xylocaine as a nerve block,’’ Brain Lang **46**, 232–244.

Herzel, H., Berry, D., Titze, I. R., and Steinecke, I. (**1995**). ‘‘Nonlinear dynamics of the voice: Signal analysis and biomechanical modelling,’’ Chaos **5**, 30–34.

Hienz, R. D., Sachs, M. B., and Sinnott, J. M. (**1981**). ‘‘Discrimination of steady-state vowels by blackbirds and pigeons,’’ J. Acoust. Soc. Am. **70**, 699–706.

Hienz, R. D., and Brady, J. V. (**1988**). ‘‘The acquisition of vowel discrimination by nonhuman primates,’’ J. Acoust. Soc. Am. **84**, 186–194.

House, A. S., and Stevens, K. N. (**1956**). ‘‘Analog studies of the nasalization of vowels,’’ J. Speech Hear Disord. **21**, 218–232.

Jackson, M. T., Espy-Wilson, C., and Boyce, S. E. (**2001**). ‘‘Verifying a vocal tract model with a closed side branch,’’ J. Acoust. Soc. Am. **109**, 2983–2987.

Kent, R. D. (**1978**). ‘‘Imitation of synthesized vowels by preschool children,’’ J. Acoust. Soc. Am. **63**, 1193–1198.

Kent, R. D. (**1979**). ‘‘Isovowel lines for the evaluation of vowel formant structure in speech disorders,’’ J. Speech Hear. Disord. **44**, 513–521.

Kojima, S., and Kiritani, S. (**1989**). ‘‘Vocal-auditory functions in the chimpanzee: Vowel perception,’’ Int. J. Primatol. **10**, 199–213.

Lieberman, P. H., Klatt, D. H., and Wilson, W. H. (**1969**). ‘‘Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates,’’ Science (Washington, DC, U.S.) **164**, 1185–1187.

Maeda, S. (**1993**). ‘‘Acoustics of nasalization and articulatory shifts in French nasal vowels,’’ in *Nasals, Nasalization, and the Velum*, edited by M. K. Huffman and R. A. Krakow, Phonetics and Phonology, Vol. 5 (Academic, San Diego), pp. 147–170.

Markel, J. D., and Gray, A. H. (**1976**). *Linear Prediction of Speech* (Springer, New York).

Mergell, P., Fitch, W. T., and Herzel, H. (**1999**). ‘‘Modeling the role of nonhuman vocal membranes in phonation,’’ J. Acoust. Soc. Am. **105**, 2020–2028.

Nowicki, S. (**1987**). ‘‘Vocal tract resonances in oscine bird sound production: Evidence from bird songs in a helium atmosphere,’’ Nature (London) **325**, 53–55.

Owren, M. J. (**1990a**). ‘‘Acoustic classification of alarm vocalization by vervet monkeys (*Cercopithecus aethiops*),’’ J. Acoust. Soc. Am. **101**, 2951–2963.

Owren, M. J. (**1990b**). ‘‘Acoustic classification of alarm vocalization by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*). I. Natural vocalization,’’ J. Comp. Psychol. **104**, 20–28.

Owren, M. J. (**1990c**). ‘‘Acoustic classification of alarm vocalization by vervet monkeys and (*Cercopithecus aethiops*) and humans (*Homo sapiens*). II. Synthetic vocalization,’’ J. Comp. Psychol. **104**, 29–40.

Owren, M. J., and Bernacki, R. H. (**1988**). ‘‘The acoustic features of vervet monkey alarm vocalization,’’ J. Acoust. Soc. Am. **83**, 1927–1935.

Owren, M. J., and Linker, C. D. (**1995**). ‘‘Some analysis techniques that may be useful to acoustic primatologists,’’ in *Current Topics in Primate Vocal Communication*, edited by E. Zimmermann, J. Newman, and U. Jürgens (Plenum, New York).

Owren, M. J., and Bernacki, R. H. (**1998**). ‘‘Applying Linear Predictive Coding (LPC) to frequency-spectrum analysis of animal acoustic signals,’’ in *Animals Acoustic Communication*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer, Berlin), pp. 129–162.

Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (**1997**). ‘‘The acoustic features of vowel-like grunt vocalization in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and function,’’ J. Acoust. Soc. Am. **101**, 2951–2963.

Peterson, G. E., and Barney, H. L. (**1952**). ‘‘Control methods used in a study of vowels,’’ J. Acoust. Soc. Am. **24**, 175–184.

Rendall, D., Owren, M. J., and Rodman, P. S. (**1998**). ‘‘The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalization,’’ J. Acoust. Soc. Am. **103**, 602–614.

Rendall, D., Seyfarth, R. M., Cheney, D. L., and Owren, M. J. (**1999**). ‘‘The meaning and function of grunt variants in baboons,’’ Anim. Behav. **57**, 583–592.

Riede, T., and Fitch, T. (**1999**). ‘‘Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*),’’ J. Exp. Biol. **202**, 2859–2867.

Riede, T., Herzel, H., Mehwald, D., Seidner, W., Trumler, E., Böhme, G., and Tembrock, G. (**2000**). ‘‘Nonlinear phenomena in the natural howling of a dog–wolf mix,’’ J. Acoust. Soc. Am. **108**, 1435–1442.

Riede, T., and Zuberbühler, K. (**2003**). ‘‘Pulse register phonation in Diana monkey alarm calls,’’ J. Acoust. Soc. Am. **113**, 2919–2926.

Robb, M. P., and Saxman, S. H. (**1988**). ‘‘Acoustic observations in young children’s non-cry vocalization,’’ J. Acoust. Soc. Am. **83**, 1876–1882.

Robb, M. P., Chen, Y. , Gilbert, H. R., and Lerman, J. W. (**2001**). ‘‘Acoustic comparison of vowel articulation in normal and reverse phonation,’’ J. Speech, Lang. Hear. Res. **44**, 118–247.

Seyfarth, R. M., and Cheney, D. L. (**2003**). ‘‘Signalers and receivers in animal communication,’’ Annu. Rev. Psychol. **54**, 145–173.

Sinnott, J. M. (**1989**). ‘‘Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus, Macaca*) and humans,’’ J. Acoust. Soc. Am. **86**, 557–565.

Sinnott, J. M., and Kreitler, N. (**1993**). ‘‘Differential sensitivity to vowel continua in Old World monkeys (*Macaca*) and humans,’’ J. Acoust. Soc. Am. **89**, 2421–2430.

Sinnott, J. M., Brown, C. H., Malik, W. T., and Kressley, R. A. (**1997**). ‘‘A multidimensional scaling analysis of vowel discrimination in humans and monkeys,’’ Percept. Psychophys. **59**, 1214–1224.

Sommers, M. S., Moody, D. B., Prosen, C. A., and Stebbins, W. C. (**1992**). ‘‘Formant frequency discrimination by Japanese macaques (*Macaca fuscata*),’’ J. Acoust. Soc. Am. **91**, 3499–3510.

Stevens, K. N., Fant, G., and Hawkins, S. (**1987**). ‘‘Some acoustical and perceptual correlates of nasal vowels,’’ in *Festschrift for Ilse Lehiste*, edited by R. Channon and L. Shockey (Foris, Dordrecht), pp. 241–245.

Story, B. H., Titze, I. R., and Hoffman, E. A. (**1996**). ‘‘Vocal tract area functions from magnetic resonance imaging,’’ J. Acoust. Soc. Am. **100**, 537–554.

Tembrock, G. (**1996**). *Akustische Kommunikation der Säugetiere* (Wissenschaftl. Buchgesell., Darmstadt).

Titze, I. R. (**1994**). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).

van den Berg, J. (**1958**). ‘‘Myoelastic-aerodynamic theory of voice production,’’ J. Speech Hear. Res. **1**, 227–244.

Wilden, I., Herzel, H., Peters, G., and Tembrock, G. (**1998**). ‘‘Subharmonics, biphonation, and deterministic chaos in mammal vocalization,’’ Bioacoustics **9**, 171–196.

Zuberbühler, K. (**2000a**). ‘‘Referential labeling in Diana monkeys,’’ Anim. Behav. **59**, 917–927.

Zuberbühler, K. (**2000b**). ‘‘Causal cognition in a nonhuman primate: Field playback experiments with Diana monkeys,’’ Cognition **76**, 195–207.

Zuberbühler, K., Cheney, D. L., and Seyfarth, R. M. (**1999**). ‘‘Conceptual semantics in a nonhuman primate,’’ J. Comp. Psychol. **113**, 33–42.

Zuberbühler, K., Jenny, D., and Bshary, R. (**1999**). ‘‘The predator deterrence function of primate alarm vocalization,’’ Ethology **105**, 477–490.