Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques

W. Tecumseh Fitch
Program in Speech and Hearing Sciences and Department of Anthropology, 33 Kirkland Street, Room 1036, Harvard University, Cambridge, Massachusetts 02138

(Received 11 December 1996; accepted for publication 14 April 1997)

Body weight, length, and vocal tract length were measured for 23 rhesus macaques (*Macaca mulatta*) of various sizes using radiographs and computer graphic techniques. Linear predictive coding analysis of tape-recorded threat vocalizations was used to determine vocal tract resonance frequencies (''formants'') for the same animals. A new acoustic variable is proposed, ''formant dispersion,'' which should theoretically depend upon vocal tract length. Formant dispersion is the averaged difference between successive formant frequencies, and was found to be closely tied to both vocal tract length and body size. Despite the common claim that voice fundamental frequency (F_0) provides an acoustic indication of body size, repeated investigations have failed to support such a relationship in many vertebrate species including humans. Formant dispersion, unlike voice pitch, is proposed to be a reliable predictor of body size in macaques, and probably many other species. © *1997 Acoustical Society of America.* [S0001-4966(97)01808-0]

PACS numbers: 43.80.Ka, 43.80.Jz, 43.70.Aj, 43.64.Bt [FD]

INTRODUCTION

The acoustic characteristics of animal vocalizations have been postulated to provide information about many important attributes of the vocalizer, including its size, age, sex, and reproductive status and emotional state. A particularly important set of parameters may relate to the size of the vocalizer. Body size is an extremely important variable in animal physiology, mortality, and ecology (Peters, 1983; Schmidt-Nielsen, 1984; Harvey, 1990; Alexander, 1996), and in many types of animal social behavior, including most prominently aggressive interactions (Parker, 1974; Clutton-Brock and Albon, 1979) and mating behavior (Clutton-Brock *et al.*, 1977). The size of various components of the sound production apparatus (e.g., the lungs, vocal folds, and vocal tract) has an important effect on the acoustic output (Fant, 1960; Lieberman, 1984). Because the size of these production components may in many cases be related to the overall weight or length of the animal, there is good reason to expect that some aspects of the acoustic signal may provide cues to the size of the vocalizer.

The most frequently cited acoustic parameter which could provide a cue to body size is mean and/or lowest fundamental frequency (Darwin, 1871). In particular, the lowest producible fundamental frequency of phonation ($F_{0\min}$) is determined by the length and mass of the vocal folds: the larger the folds, the lower is $F_{0\min}$. If the size of the vocal folds were related to the vocalizer's body size, $F_{0\min}$ would provide a good cue to body size (Morton, 1977; Hauser, 1993). This indeed appears to be the case in some species, including some toads and frogs (Martin, 1972; Davies and Halliday, 1978; Ryan, 1988). However, such a relationship between body size and vocal fold size does not seem to be typical in other vertebrates. For instance, there is no correlation between F_0 and body size in adult humans (Lass and Brown, 1978; Künzel, 1989; Cohen *et al.*, 1980; van Dommellen, 1993), red deer (McComb, 1991), and other amphibian species (Asquith and Altig, 1990; Sullivan, 1984). This lack of correlation in adult humans is particularly surprising given the widespread assumption that a ''deep'' or low-pitched voice indicates large body size.

The lack of correlation between F_0 and size seems less surprising when the anatomy of the vocal folds is considered. The folds are housed within the flexible cartilaginous larynx, which itself floats at the top of a trachea and is unconstrained in size by neighboring bony structures [the hyoid bone, though ossified, grows as a unit with the larynx, Schön (1971) and Schneider *et al.* (1967)]. Thus the larynx and vocal folds can grow independently of the rest of the head or body, as indeed occurs in human males at puberty (Negus, 1949; Goldstein, 1980). At puberty, androgen receptors in the laryngeal cartilages respond to increased circulating testosterone with a profound growth spurt (Tuohimaa *et al.*, 1981; Beckford *et al.*, 1985). The result is a typical F_0 for adult males which is about half that of adult females, despite an average difference in body weights of only 20% (Hollien, 1960). Hypertrophy of the male larynx, out of all proportion to body size, is carried to an absurd extreme in animals such as the howler monkey [*Allouatta seniculus*, Schön (1971)], in which the larynx and hyoid together are the size of the entire skull, or the hammerhead bat [*Hypsignathus monstrosus*, Kingdon (1974) and Schneider *et al.* (1967)], where the larynx of the male occupies virtually the entire thoracic cavity.

When such developmental flexibility is present there is clearly no *a priori* reason to expect vocal fold size (and thus F_0) to be well-correlated with body size (Fitch, 1994; Fitch and Hauser, 1995). Despite the common claim that voice pitch provides an accurate cue to body size (e.g., Morton, 1977), these data suggest that the larynx is ill-suited to provide dependable cues to body size.

A different potential acoustic cue to body size derives from the fact that, in most vertebrates, the sound signal created in the larynx passes through the supralaryngeal vocal tract (hereafter, simply ''vocal tract'') before being radiated into the environment. The column of air in the vocal tract has certain natural modes of vibration or formants, which affect the resultant output signal (Fant, 1960) (the terminology of speech scientists is adopted here, using ''formants'' as opposed to the more general term ''resonances,'' despite the fact that the sounds produced by nonhumans differ in significant ways from human speech). If the cross-sectional area function of the vocal tract is constant, the primary determinant of formant frequencies is the length of the vocal tract (Fant, 1960; Lieberman and Blumstein, 1988). In particular, a lengthening of the vocal tract tube will lead to a decrease in the average spacing between successive formants, or ''formant dispersion.''

Several researchers in speech science have postulated a relationship between body size and formant frequencies (Fant, 1960; Nearey, 1978; Lieberman, 1984; Peterson and Barney, 1952), and Peterson and Barney provided evidence suggestive of such a relationship in humans (children have higher formant dispersion than adults, and women greater than men). Because the mammalian vocal tract is made up of the pharyngeal, oral, and nasal cavities, which are firmly bounded by the bones of the skull, and skull size is closely tied to overall body size (Morita and Ohtsuki, 1973; Dechow, 1983; Janis, 1990; Valkenburgh, 1990; Alcantara *et al.*, 1991; Sharma, 1990; and see data below), vocal tract length should be much less free to vary independently of body size than larynx size. Thus we can expect vocal tract length and the attendant acoustic cue of formant dispersion to provide a correspondingly more robust cue to body size.

Systematic investigation of the role of the supralaryngeal vocal tract in nonhuman vocalization was inaugurated with the work of Lieberman (Lieberman, 1968; Lieberman *et al.*, 1969; Lieberman, 1975; Lieberman, 1984), who was interested primarily in the differences between the human and nonhuman vocal tract. Lieberman showed that production of the full range of vowels, in particular the /i/ and /u/ vowels which are ubiquitous in human language, is impossible without the ''two-tube'' vocal tract of modern human beings. However, Lieberman (1968, p. 1576) also described modifications of monkey calls via changes in vocal tract length, and did not rule out the possibility that such changes have communicative significance. Further suggestions of a possible communicative role for vocal tract resonances in nonhumans came from spectrographic analyses of baboons (*Papio hamadryas*), geladas (*Theropithecus gelada*), and vervets (*Cercopithecus aethiops*) (Andrew, 1976; Richman, 1976; Seyfarth and Cheney, 1984, respectively).

More recent work has documented the role of supralaryngeal articulation in modifying acoustic characteristics of nonhuman vocalizations. Bauer (1987) built on Marler and Tenaza's (1977) technique of frame-by-frame video analysis of vocal production in chimpanzees (*Pan troglodytes*) to show that F_0 is related to mouth-opening size during chimp vocalizations. However, whether any causal acoustic mechanism underlies this correlation remains unclear. Hauser *et al.* (1993) used video analysis to analyze rhesus macaque calls, showing that changes in the first resonance frequency were well-correlated with changes in mandible position. A similar result was documented in cat (*Felis domesticus*) vocalizations by Shipley *et al.* (1991). Finally, Hauser and Schön-Ybarra (1994) experimentally eliminated vocal tract elongation via lip movements (using injections of the nerve-blocker xylocaine into the peri-oral region of rhesus macaques). They found that resonance frequencies were significantly higher than normal in the ''coo'' vocalization, which is normally accompanied by rounded lips. No changes were observed in F_0 or call duration. These experimental data are thus consistent with the predictions of source–filter theory (Fant, 1960), and indicate independence of source and filter in these calls.

Overall, a wealth of data suggests that the principles of source–filter theory and acoustic phonetics, originally developed for human speech, are applicable to nonhuman vocalizations as well. Source and filter appear to be independent in most cases, and movements of the articulators affect the vocal tract filter in the predicted ways. A variety of vertebrates can use the differences in formant frequency to discriminate synthesized vowels (baboons: Heinz and Brady, 1988; dogs: Baru, 1975; cats: Dewson, 1964; blackbirds and pigeons: Heinz *et al.*, 1981), and macaques can perceive formants with accuracy rivaling that of humans (Sommers *et al.*, 1992). However, the only conclusive evidence that supralaryngeal acoustic cues are utilized in nonhuman communication comes from Owren's work with vervet monkeys (*Cercopithecus aethiops*). Owren and Bernacki (1988) used linear predictive coding (LPC) analysis of vervet ''snake'' and ''eagle'' alarm calls to isolate characteristics of source waveform, presumed vocal tract filtering functions, and temporal patterning which distinguished these calls. Owren (1990b) then used an operant paradigm to test classification of synthetic calls in which each of these characteristics was modified independently. The results indicated that spectral characteristics played the dominant perceptual role in distinguishing the two call types, suggesting that supralaryngeal articulation conveys distinctive information in these calls.

The literature reviewed above suggests that vocal tract length and formant frequencies may be correlated with body size in many vertebrate species, and that nonhumans possess the perceptual machinery to make use of this correlation. However, despite the fact that several researchers have suggested such a relationship (Fant, 1960; Lieberman, 1984; Peterson and Barney, 1952), no study has addressed this issue directly in any species. The goal of this study was to measure body weight and body length, the lengths of the oral and nasal vocal tracts, and formant frequency dispersion in calls, and to quantify the relationship(s) between these variables, using rhesus macaques as the study species.

I. METHODS

A. Study animals

The subjects were 20 healthy rhesus macaques (*Macaca mulatta*), housed at the New England Regional Primate Research Center in Southboro, MA. Animals were between 1 and 9 years of age (mean 4.7 yr), and between 2.6 and 15.6

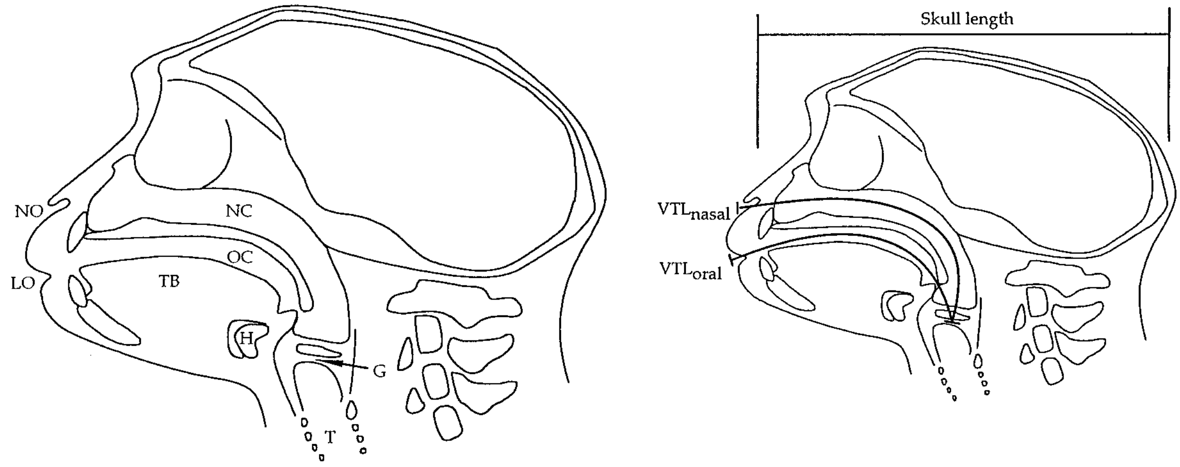

FIG. 1. Schematic of anatomical features (left) and morphometric features (right) used in this study: G: glottis, H: hyoid bone with subhyoid air sac, OC: oral cavity, LO: lip opening, NC: nasal cavity, NO: nostril opening, T: tracheal lumen, TB: tongue body, VTL: vocal tract length.

kg in weight (mean 7.13 kg). Both males ($N=11$) and females ($N=9$) were studied. Animals were fed primate chow, fresh fruit, and *ad libitum* water, and were maintained at their *ad libitum* body weight.

B. Anatomical measures

Animals were anesthetized with 5–7 mg/kg of Telazol (a standard veterinary anesthetic mixture of tiletamine and zolazepam), or 10 mg/kg of Ketaset (ketamine hydrochloride), rendering them unconscious for approximately 1/2 h. They were weighed on a Mettler Toledo SM 30-K digital electronic balance to an accuracy of ±10 g, and their crown–rump length was measured to centimeter accuracy. Each animal was laid upon its side on the radiographic table, and midsagittal radiographs (hereafter x rays) were made of the head and neck region (1/40-s exposure time, 100 mA, 92–116 kV depending on animal size). To allow accurate determination of absolute size from the x-ray films, a 3-cm lead reference strip was placed within the area of exposure, held above the table by a cardboard support at the animals' midsagittal height.

X-ray clarity was sufficient to delineate the outlines of both the oral and nasal vocal tracts, as well as the location of the glottis (Fig. 1). The location of the glottis was made clearer by the presence of a subhyoid air sac in this species (Geist, 1933), which opens into the larynx via a thin tube directly above the level of the ventricular folds, and was clearly visible in the x-ray images.

Vocal tract length (VTL) was determined from tracings of the x-ray images using a Wacom ArtZ UD 0608 digitizing tablet and NIH Image software (version 1.52, available free on the Internet from NIH). A curvilinear line equidistant from the medial and external walls of the oral or nasal vocal tract was drawn from the middle of the glottis to the opening of the lips or external nares, respectively (Fig. 1). This measure of VTL is acoustically motivated: It follows the path of plane-wave propagation of sound from the glottis to the oral or nasal radiation site, and should on theoretical grounds be associated with formant frequency dispersion (Fant, 1960). The VTL, in pixels, was recorded along with the length of the 3-cm lead strip, in pixels, which was later used to derive the actual VTL in cm. The length of the entire skull from the occipital ridge to the front of the incisors was also recorded.

The accuracy of this technique was very high: ten repeated measures of one animal's x ray yielded standard deviations of 5 pixels or less (1 mm or less). Standard errors were a fraction of a millimeter for all x-ray measures.

C. Acoustic measures

Recordings were made with a Sony WM-D3 Walkman Professional recorder and Sony PC-62 microphone using Maxell XL II high-bias cassette tapes. Animals were recorded in their home cages before being captured or anesthetized. Lip configuration was carefully observed during recording, because macaques, like humans, can substantially shorten their vocal tract by retracting the lips, or lengthen it by protruding the lips (Hauser and Schön-Ybarra, 1994). Animals typically emitted aggressive ''pant-threats'' and ''pant-barks'' (Hauser *et al.*, 1993) when faced or stared at by an unfamiliar observer. These vocalizations were made with an almost completely closed mouth and no lip protrusion (the ''tense-mouth face'' of van Hoof, 1967), and thus should be closely registered with the anatomical measure of oral vocal tract length employed in this study. Threat calls are brief, noisy, coughlike calls (Fig. 2 and Bercovitch *et al.*, 1995), with a very broadband source, making them well-suited for formant frequency measurement (unlike a high-frequency tonal call, where formants are only detectable if a harmonic and formant coincide, see Lieberman and Blumstein, 1988; Ryalls and Lieberman, 1982). Therefore although some other call types were recorded (mainly coos), all acoustic analyses were performed on these threat vocalizations.

Recordings were digitized at 16 bits quantization and 22-kHz sampling rate using the built-in sound input on an Apple Power Macintosh 6100/60, using Macromedia SoundEdit 16 software. Formant frequencies were measured using LPC analysis (Markel and Gray, 1976; Wakita, 1976), implemented via autocorrelation in Matlab 4.2 (using the Matlab Signal Processing Toolbox). LPC analysis has been

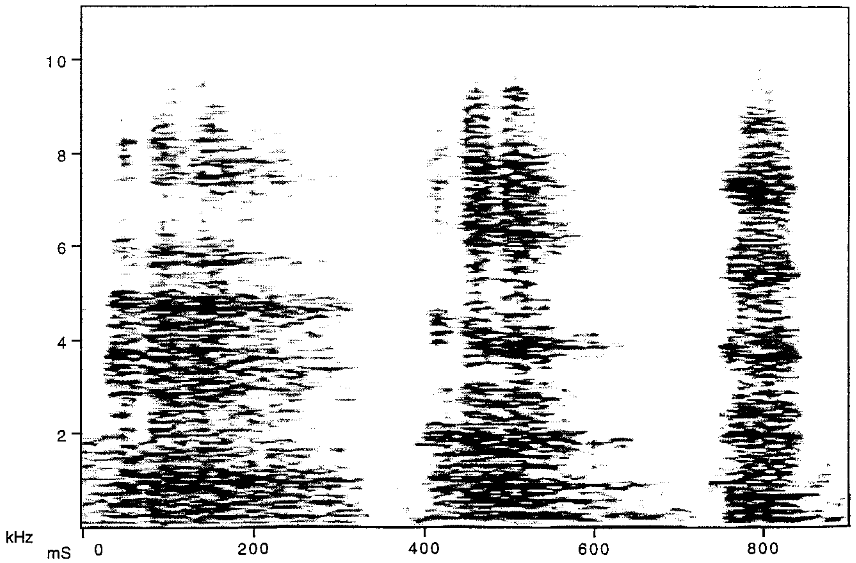

FIG. 2. Spectrograms of threat vocalizations from three different individual *M. mulatta* of increasing size. Note the decreasing formant dispersion (from left to right: 2.6 kg female; 5.3 kg female; 9.2 kg female) (2048-point Hamming window FFT, 50% frame overlap, 176 Hz bandwidth).

used to analyze similar primate vocalizations (Owren and Bernacki, 1988; Rendall, 1996), and its applicability has been verified by Owren (1990a, b).

Vocal tract length variation between monkeys should theoretically result in a variable number of formants below the Nyquist frequency, with short vocal tracts producing few formants and long tracts producing many. Using a simple quarter-resonator tube model (Lieberman and Blumstein, 1988), a 5-cm vocal tract would have three formants below our Nyquist frequency of 11 kHz, while a 10-cm vocal tract would have seven. A variable number of poles was used in the LPC analysis (3–9; LPC order=8–20). All LPC measurements were visually verified by superimposing the LPC-derived frequency response over an 512-point fast Fourier transform (FFT) of the same time slice, allowing the user to empirically determine the optimum number of poles for each call by trial and error.

The 512-sample arrays from locations chosen interactively by the user were input to the LPC function (no pre-emphasis, no weighting window). At least five LPC spectra were derived by analyzing successive time-slices of each call. Due to the noisy and random nature of the glottal source in pant-threats, all formants were not excited in each time slice; therefore, the successive LPC spectra were averaged together resulting in a single long-term averaged LPC spectrum. Typically, formant locations did not change appreciably during the course of a call; occasionally there was slight (±100 Hz) variation in a given formant frequency over the course of the call (probably due to slight mouth closing or opening).

Animals varied greatly in the number of usable vocalizations produced, with some producing only a few and others producing dozens. However, due to noise in the animal room from cage movements and other animals, most call recordings obtained were unanalyzable. Fortunately, formant frequency measurements are extremely consistent from one call token to the next (see Results). Thus for most analyses, the single best call (highest signal-to-noise ratio) was chosen (if there were multiple possibilities, the first one on the tape was chosen). A single call was used for each animal because for many animals only a single good-quality call was available, and this avoided the statistical complications caused by mixing single measurements for some animals with mean values for others.

Formant dispersion (D_f), which is the average distance between each adjacent pair of formants, was calculated using the following formula:

$$D_f = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N-1}, \tag{1}$$

where D_f is the formant dispersion (in Hz), N is the total number of formants measured, and F_i is the frequency (in Hz) of formant i.

II. RESULTS

Summary data for each variable measured are given in Table I, and a summary of all the correlations examined in this study is given in Table II. In general, intercorrelations were very strong between vocal tract length (VTL), formant dispersion, and both measures of body size. All of the variables measured were roughly normally distributed, so the use of parametric statistics (i.e., regression analysis) was justified.

Although males tended to be slightly larger than females, there was no significant sex difference in any anatomical or acoustic variable measured (unpaired t tests, $p > 0.05$). Therefore, unless otherwise noted all of the analyses reported here combine data from males and females.

A. Anatomical correlations

Extremely tight correlations were found among measurements of body size, skull length, and vocal tract length. Body weight and body length were highly correlated

TABLE I. Basic descriptive data for acoustic and anatomical variables. (σ is the standard deviation, S.E. the standard error of the mean, N the count, and ''min'' and ''max'' the minimum and maximum values).

	Mean	σ	S.E.	N	Min.	Max.	# Missing
Age (yr)	4.667	2.590	0.610	18	1.000	9.000	2
Weight (kg)	7.128	3.774	0.844	20	2.600	15.580	0
Body length (cm)	47.417	6.576	1.550	18	35.000	60.000	2
Skull length (cm)	12.204	1.558	0.348	20	9.254	15.022	0
Oral VTL (cm)	7.850	1.290	0.289	20	5.514	9.739	0
Nasal VTL (cm)	8.915	1.348	0.301	20	6.739	10.891	0
# Formants	4.800	1.399	0.313	20	3.000	7.000	0
Formant D (kHz)	2.146	0.700	0.157	20	1.250	3.510	0
Log_{10}wt	0.792	0.243	0.054	20	0.415	1.193	0

($r=0.911$, $p<0.0001$), as were body length and $\log_{10}$ body weight ($r=0.953$, $p<0.0001$) (Fig. 3). Because body weight should theoretically be proportional to the cube of a linear dimension, $\log_{10}$ body weight is used hereafter (logarithms were not taken for all measurements because the relationships between body length, skull length, VTL, and formant dispersion appear to be linear).

Skull length was correlated with body length ($r=0.974$, $p<0.0001$) and $\log_{10}$ weight ($r=0.955$, $p<0.0001$). VTL was correlated with skull length (oral VTL, $r=0.957$, $p<0.0001$; nasal VTL, $r=0.927$, $p<0.0001$), as was expected given that the main dimensions of the vocal tract are delineated by the skull (Fig. 3). As expected due to the intercorrelations between body size, skull length, and VTL, VTL correlated with $\log_{10}$ body weight (oral VTL, $r=0.947$, $p<0.0001$; nasal VTL, $r=0.906$, $p<0.0001$) and body length (oral VTL, $r=0.950$, $p<0.0001$, Fig. 3; nasal VTL, $r=0.928$, $p<0.0001$).

B. Acoustic correlations

The number of usable pant-threat calls obtained varied greatly among animals (range 1–10, mode=1, for four animals). For four animals, seven or more calls were obtained, allowing an analysis of consistency in formant locations among the calls of a particular individual. The formant values for these monkeys are shown in Fig. 4, which illustrates the impressive consistency of formant locations for each individual, and the substantial differences between individuals. Therefore, the rest of the analyses were performed with only one call per animal, as explained in the Methods section.

There was variation in the number of formants found for different animals. The number of formants below the 11 kHz Nyquist frequency (''formant density'') varied from three to seven (mean 4.8), and as predicted by theory was correlated with oral vocal tract length ($r=0.828$, $p<0.0001$)and inversely correlated with formant dispersion ($r=-0.918$, $p<0.0001$). Because formant density is an integer value, it is less sensitive than formant dispersion and is not further used (see Sec. III).

There were always at least three formants for any given call. Therefore, acoustic analyses were done in two ways for each call: including all formants present (thus making use of all information available for any given call), or including only the lowest three formants per call (assuring equality of analysis between animals). The pattern of results was identical for the two analyses, and therefore only the measure of formant dispersion defined in Eq. (1), which includes all available formats, is reported.

Formant dispersion was correlated significantly and negatively with VTL, as predicted by the source–filter theory of vocal production (oral VTL, $r=-0.915$, $p<0.0001$; nasal VTL, $r=-0.852$, $p=0.0001$). Given the anatomical correlations described above, it is thus unsurprising that formant dispersion correlated with $\log_{10}$ body weight ($r=-0.868$, $p<0.0001$) and body length ($r=-0.816$, $p<0.0001$). Stepwise multiple regression analysis indicated that $\log_{10}$ body weight accounted for the most variance in formant dispersion (partial correlation coefficients: age -0.713, body length -0.821, $\log_{10}$ body weight -0.876), and that it alone accounted for all significant variance in formant dispersion.

The formant dispersion (D_{pred}, in Hz) predicted by a simple one-tube model of the vocal tract (without end correction, since no data was available on size of the oral opening) is

TABLE II. Correlation coefficients between the various acoustic and anatomical variables measured in this study. All correlations are significant at the $p<0.0001$ level.

	1	2	3	4	5	6	7
1. Log_{10} wt	1.000	0.942	0.944	0.904	−0.723	−0.886	0.953
2. Skull length	0.942	1.000	0.963	0.939	−0.711	−0.869	0.974
3. Oral VTL	0.944	0.963	1.000	0.972	−0.724	−0.922	0.950
4. Nasal VTL	0.904	0.939	0.972	1.000	−0.625	−0.868	0.928
5. $F1$	−0.723	−0.711	−0.724	−0.625	1.000	0.699	−0.702
6. Formant D	−0.886	−0.869	−0.922	−0.868	0.699	1.000	−0.816
7. Body length	0.953	0.974	0.950	0.928	−0.702	−0.816	1.000

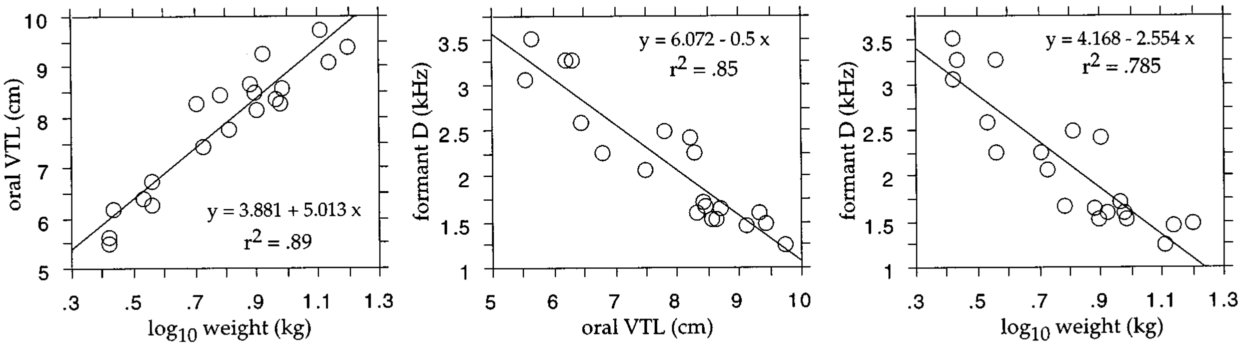

FIG. 3. Bivariate plots illustrating intercorrelations of the base 10 logarithm of weight (kg), oral vocal tract length (cm), and formant dispersion (kHz).

$$D_{\text{pred}}=\frac{c}{2L}, \tag{2}$$

where c is the speed of sound (335 m/s) and L is the vocal tract length in m. This prediction is the same for open or closed tubes (where the resonances are related as 2,4,6,...), or half-open tubes (where the relation is 1,3,5,...). The formant dispersion observed did not differ significantly from that predicted by Eq. (2) (paired t tests, $t=0.63$, $p=0.53$). Thus although the x-ray data indicate that the monkey vocal tract is not a tube of exactly uniform diameter, use of a more complex model of vocal tract anatomy than a simple tube (as suggested by Shipley *et al.*, 1991) does not seem warranted by the current data. The good fit between predicted and actual values also provides further evidence of the usefulness of low-order LPC for isolating and measuring spectral peaks of nonhuman vocal tracts (Owren and Bernacki, 1988; Owren, 1990a,b). Because the excitation signal for the threat vocalizations analyzed here is coughlike and noisy (presumably generated by noise at the glottis), it is plausible to interpret the low-order LPC spectrum as characterizing the vocal tract filter function, and thus the spectral peaks in the LPC spectrum as formant frequencies.

These results were consistent across both ages and sexes. When only adult animals were included in the analysis (i.e., animals aged 4 years or older), the correlation between $\log_{10}$ body weight and formant dispersion dropped (due to the smaller sample size) but was still very significant statistically ($N=11$, $r=-0.73$, $p=0.008$). The correlation also held for juveniles only ($N=7$, $r=-0.84$, $p=0.015$). Similarly, the correlation between $\log_{10}$ body weight and formant dispersion was strong in the 11 males considered separately ($r=-0.896$, $p=0.0002$), as in the nine females separately ($r=-0.913$, $p=0.0006$). Thus the relationship between body size and acoustic output described here does not result simply from differences between adults and juveniles, or between males and females. Instead, it appears to result from a direct anatomical relationship between body size, vocal tract length, and formant dispersion.

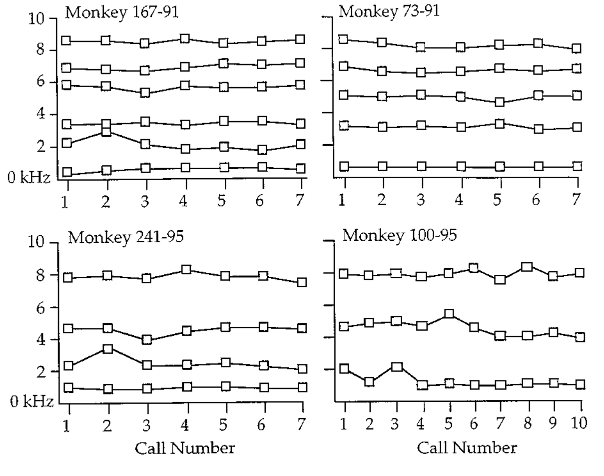

FIG. 4. Consistency of formant frequency measurements across different calls for each of four individual monkeys.

III. DISCUSSION

The results reported here indicate that formant dispersion, which is determined by vocal tract length (VTL), is closely correlated with body weight and length in rhesus macaques. As predicted by source–filter acoustic theory, in tandem with the tight anatomical correlations reported above, the spacing between formants provides an accurate means of predicting a vocalizer's body mass and/or length. Formant dispersion in the threat vocalizations of rhesus macaques thus provides an excellent *potential* cue to body size in the population of rhesus macaques studied here (r^2 greater than 75%). Whether rhesus macaques use this information remains unknown, but a growing body of data (reviewed in the Introduction) strongly suggest that nonhuman animals attend to formant cues and make use of them in their species-specific communication systems (Dewson, 1964; Lieberman, 1968; Baru, 1975; Andrew, 1976; Richman, 1976; Heinz *et al.*, 1981; Seyfarth and Cheney, 1984; Heinz and Brady, 1988; Owren and Bernacki, 1988; Owren, 1990a,b; Sommers *et al.*, 1992; Hauser and Schön-Ybarra, 1994; Hauser *et al.*, 1993; Fitch, 1994; Fitch and Hauser, 1995; Rendall; 1996; Owren *et al.*, 1997). Rendall (1996) suggested that one use of formant information in macaque communication is to provide a relatively stable indicator of individual identity. The data in this paper support this idea (because formants were found to be very stable over multiple calls), and suggest that another potential use for formant information is to evaluate a vocalizer's body size.

A. Reliability of formant dispersion

Despite the common conception that F_0 provides a cue to body size (e.g., Morton, 1977), the preponderance of evidence now suggests that F_0 does not provide an accurate indication of size in adult humans (Lass and Brown, 1978; Künzel, 1989; Cohen *et al.*, 1980; van Dommellen, 1993) or in several other animals (McComb, 1991; Asquith and Altig, 1990; Sullivan, 1984). I suggest that this lack of correlation is due to the relative independence between body size and larynx size (which determines vocal fold length and hence lowest F_0). The cartilages of the larynx can grow independently of the rest of the body, and experience a hormone-dependent growth spurt in many mammals (Tuohimaa *et al.*, 1981; Beckford *et al.*, 1985; Hollien, 1960) that leads to a decoupling of body size and F_0, particularly in males. In contrast, vocal tract length is anatomically dependent upon skull size, which is in turn closely correlated with body size in all mammal species which have been examined (Morita and Ohtsuki, 1973; Dechow, 1983; Janis, 1990; Valkenburgh, 1990; Alcantara *et al.*, 1991; Sharma, 1990). The results of the current study indicate that (a) body size, skull size, and VTL are anatomically correlated in macaques, and (b) VTL determines formant dispersion, which thus provides an acoustic indication of body weight and mass. Hence, unlike F_0, formant dispersion should provide a robust, reliable cue to body size in most mammals and perhaps terrestrial vertebrates in general.

Two key factors underlie this reliability: the link between formant dispersion and VTL, and the link between VTL and body size. Each of these are discussed in turn. Because tract length affects the overall pattern of formant frequencies, any given formant provides some information regarding tract length. Thus, for example, the lowest formant ($F1$) could potentially be used as a cue to tract length. However, using one formant as the sole cue to VTL would entail several problems. First, if the frequency band of the chosen formant is obscured by environmental noise or degradation, no VTL information would be available. In contrast, formant dispersion capitalizes on the redundancy of the formant spacing pattern, and is thus robust to the degradation of information in any one (or even several) formants. Second, individual formants are more sensitive to changes in the vocal tract transfer function (as seen in human vowels and, to a lesser degree, in baboon grunts: Owren *et al.*, 1997). As a statistical measure encompassing all formant information, formant dispersion is less sensitive to deviations in a single formant.

Another potential cue to VTL is ''formant density'' (the number of formants in a particular frequency range, e.g., in this study the total number of formants below the 11-kHz Nyquist frequency), which is inversely related to formant dispersion. The main problem with density as an acoustic cue to tract length is that, as an integer measure, it is considerably less sensitive to changes in VTL than formant dispersion, which is as accurate as discrimination of individual formants [formant difference limens are 1%–5% in both humans and macaques, Sommers *et al.* (1992)]. Of course, it would be possible to devise more accurate formant density indices, but the simplicity and numerical tractability of formant dispersion recommend it. Thus formant dispersion is accurate, robust to environmental deterioration, and resistant to errors due to variability in individual formants.

B. Source of unreliability for vocal tract cues

When it comes to providing vocal tract information, not all calls are equal. If the source contains energy at only a few, widely spaced frequencies (e.g., a high-pitched tonal call), it will provide little information about the vocal tract transfer function. Thus low-pitched calls provide for better resolution of formants than do sounds with high F_0's (Ryalls and Lieberman, 1982). Better yet, a noisy or impulsive source is ideally suited for accurately outlining the transfer function (e.g., Tartter and Braun, 1994 with whispers). The accuracy of formant dispersion as an acoustic cue to body size will thus depend on the particular type of call, with harsh, noisy or impulsive calls (like the threat vocalizations studied here) being ideal, and high-pitched tonal calls being worst. As well as having clear methodological implications, this observation leads to an interesting behavioral prediction. In contexts where accurate information about size is favored (e.g., a large animal demonstrating its size to competitors or mates), noisy or low-pitched calls might be expected since they provide more accurate delineation of vocal tract cues (see also Peters, 1984). This provides alternative grounds for Morton's (1977) ''motivational-structural rule'' that low, noisy sounds accompany aggression.

Another potential source of unreliability is changes in vocal tract length due to articulatory movements. Opening or nearly closing the mouth, or (in animals with mobile lips) protruding or retracting the lips will alter tract length to some degree (around 20% in humans, Fant, 1960), as will raising or lowering the larynx. However, the manipulation of VTL via lip movements does not necessarily render it (or related acoustic cues) unreliable. There is still a maximum VTL, fixed by skull dimensions, which is attained when the mouth is nearly closed and the lips are pursed. This articulatory position is characteristic of the threat vocalizations studied here, along with threats in a wide variety of other species (see below). The observed correlation between lip rounding and threat raises an interesting possibility. If listeners associate long vocal tracts with larger bodies, an animal that elongates its vocal tract while vocalizing would maximize its perceived body size. This suggests that aggressive vocalizations made with nearly closed mouths and protruded lips would be more threatening, and submissive vocalizations made with lips retraced more appeasing, and that the evolutionary origin of certain facial expressions is not purely visual, but arises also from their acoustic effects (Ohala, 1980; Fridlund, 1994; Fitch, 1994; Fitch and Hauser, 1995).

Vocal tract elongation is associated with threat behavior in many species. van Hooff (1967; p. 18), in his general review of primate facial displays, cited the ''tense-mouth face'' in which the ''mouth corners are brought forward... as a result the mouth often looks like a narrow slit.'' This display is usually performed by a dominant animal immediately preceding an attack, and is associated with a low-pitched bark in at least some species (chimpanzees and baboons). Similar ''pouts'' or ''pucker'' behavior is associated with threat in other primates (Kaufman, 1967; Fitch, 1994; Epple, 1967), seals (Miller, 1975), bears (Pruitt and Burghardt, 1977), and in dogs and other canids (Fox and Cohen, 1977).

Similarly, the use of mouth corner retraction to signal submission and absence of threat appears to be extremely widespread in mammals. A vast number of species, including dogs, seals, ungulates, and almost all primates share a submissive ''grin'' facial display (van Hooff, 1967, 1972; Miller; 1975; Fox and Cohen, 1977; Oppenheimer, 1977; Marler and Tenaza, 1977; Gautier and Gautier, 1977; Walther, 1977, 1984; Goodall, 1986; Preuschoft, 1992, 1995). This expression is variously ascribed a defensive or submissive role and is occasionally seen during play. The human smile appears to be an example of this display, although its use has been extended into nonaggressive situations to denote an unthreatening or friendly attitude (van Hooff, 1972; Preuschoft, 1992, 1995). These observations prompted Ohala (1980) to suggest that shortening the vocal tract in order to seem smaller would provide an acoustic (rather than visual) function for the human smile. The data in this paper are consistent with this hypothesis.

Finally, the anatomical correlation between VTL and skull size reported here and suspected in most mammals does not hold for all vertebrates. In birds, the sound-producing source is the syrinx, which lies at the base of the trachea near the lungs (Nowicki and Marler, 1988). Thus the trachea is an integral part of the bird vocal tract, and the length of the trachea must be added to the length of the oral/nasal cavities to derive the bird's total vocal tract length, which is not limited by the size of the skull. The trachea (like the larynx) is a flexible structure, floating free from any rigid skeletal attachments. It is thus interesting that a wide variety of birds exhibit tracheal elongation where the trachea is looped in great coils within the chest or sternum (Berndt, 1938; Neimeier, 1979). This bizarre character is found only in large birds, and often only in males, suggesting that it may represent an adaptation for acoustically exaggerating body size (Fitch, in preparation, 1994).

A second potential exception is anatomically modern *Homo sapiens*. Our species has abandoned the structural pattern typical of mammalian vocal tract anatomy. In humans the larynx has descended from its normal position at the back of the mouth to a position deeper in the throat (Negus, 1949; Laitman and Crelin, 1976; Leiberman, 1984; Flügel and Rohen, 1991), which allows the tongue body to move freely back and forth, thus creating a wider variety of vocal tract area functions than are possible in nonhumans (Lieberman, 1968; Lieberman *et al.*, 1969). Though typically viewed as an adaptation for articulate speech (Lieberman, 1975, 1984), this speech advantage would only be gained with considerable laryngeal descent (several cm). Presuming that this anatomical reconfiguration proceeded gradually over evolutionary time, what selective advantage was provided by the descent of the larynx before the conditions for improved speech were met?

The descent of the larynx from the standard mammalian position has the effect of elongating the vocal tract, possibly freeing human tract length from the skeletal size constraints described above for most mammals. Since vocal tract length is used by human listeners as a cue to body size (Fitch, 1994), it is possible that the original function of the descending larynx in early hominids was to exaggerate body size. This idea gains support from the observation that there is sexual dimorphism in the degree to which the human larynx descends (a full vertebra lower in males than females, Senecail, 1979; Harrison, 1995), with no accompanying advantage for males' vowel clarity over that of females. The degree to which vocal tract length, formant dispersion, and body size are related in humans remains an interesting matter for further research.

C. Vocal tract cues and human speech perception

One key phenomenon in human speech perception that may be related to the data in this paper is vocal tract normalization, which is the use of overall formant pattern to ''normalize'' the vowels of a given speaker (Ladefoged and Broadbent, 1957; Nearey, 1978; Lieberman, 1984). Because of variations in vocal tract length, the $F1$-$F2$ vowel space of a child is quite different from that of an adult (Peterson and Barney, 1952). A result is that children imitating adults, or adults listening to children, need to adjust their perception of a speaker's vowels to their VTL-related formant space. If a mechanism for estimating body size from formant dispersion existed in our prelinguistic ancestors, this could have provided a preadaptive basis for vocal tract normalization in humans.

The current results joins a collection of recent studies suggesting that the role of the supralaryngeal vocal tract in nonhuman animal communication may be more prominent than traditionally realized (e.g., Bauer, 1987; Owren, 1990a,b; Owren *et al.*, 1997; Shipley *et al.*, 1991; Hauser and Schön-Ybarra, 1994; Owren and Bernacki, 1988; Rendall, 1996). Although formant frequency perception plays a fundamental role in human speech perception (Joos, 1948; Lieberman and Blumstein, 1988), the evolutionary basis for this is poorly understood, mainly because we have so little information about the uses to which formant frequencies may be put in animal communication systems. The results of this work suggest that the use of formant frequency patterns to gauge body size could have been a factor promoting the original evolution of formant perception, a capability later put to such extensive and sophisticated use in human language.

ACKNOWLEDGMENTS

Many thanks are due to David Lee-Parritz for the use of the facilities at the New England Primate Center. I am very grateful to Dr. Lee-Parritz, Barbara Brewer, Virginia Rainville, and Nirah Shomer for their invaluable help in gathering data. Miguel Schön-Ybarra provided assistance in interpreting the x-ray data. Comments by Marc Hauser, Caroline Henton, Nelson Kiang, Daniel Lieberman, Philip Lieberman, Michael Owren, Ken Stevens, and an anonymous reviewer are gratefully acknowledged. This work was supported by an NIH Postdoctoral Fellowship to the author.

Alcantara, M., Diaz, M., and Pulido, F. J. P. (**1991**). ''Variabilidad en las relaciones alométricas entre el peso y las medidas craneales en el raton de campo *Apodemus sylvaticus* L. Efectos sobre su utilidad en estudios de ecologia trofica de aves rapaces,'' Doñana, Acta Vertebrata **18**, 205–216.

Alexander, R. M. (**1996**). ''Biophysical problems of small size in vertebrates,'' Symp. Zoological Soc. London **69**, 3–14.

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

Asquith, A., and Altig, R. (**1990**). ''Male call frequency as a criterion for female choice in *Hyla cinerea*,'' J. Herpetology **24**, 198–201.

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

Bauer, H. R. (**1987**). ''Frequency code: orofacial correlates of fundamental frequency,'' Phonetica **44**, 173–191.

Beckford, N. S., Rood, S. R., and Schaid, D. (**1985**). ''Androgen stimulation and laryngeal development,'' Ann. Otol. Rhinol. Laryngol. **94,** 634–640.

Bercovitch, F. B., Hauser, M. D., and Jones, J. H. (**1995**). ''The endocrine stress response and alarm vocalizations in rhesus macaques,'' Animal Beh. **49,** 1703–1706.

Berndt, R. (**1938**). ''Intrasternale Trachealschlingen bei Vogëln,'' Gegenbauer's Morphologische Jahrbuch **82**, 27–117.

Clutton-Brock, T. H., and Albon, S. D. (**1979**). ''The roaring of red deer and the evolution of honest advertising,'' Behaviour **69**, 145–170.

Clutton-Brock, T. H., Harvey, P. H., and Rudder, B. (**1977**). ''Sexual dimorphism, socionomic sex ratio and body weight in primates,'' Nature (London) **269,** 797–800.

Cohen, J. R., Crystal, T. H., House, A. S., and Neuburg, E. P. (**1980**). ''Weighty voices and shaky evidence: a critique,'' J. Acoust. Soc. Am. **68,** 1884–1886.

Darwin, C. (**1871**). *The Descent of Man and Selection in Relation to Sex* (Murray, London).

Davies, N. B., and Halliday, T. R. (**1978**). ''Deep croaks and fighting assessment in toads,'' Nature (London) **274,** 683–685.

Dechow, P. C. (**1983**). ''Estimation of body weights from craniometric variables in baboons,'' Am. J. Phys. Anthropol. **60**, 113–123.

Dewson, J. H. (**1964**). ''Speech sound discrimination by cats,'' Science **141,** 555–556.

Epple, G. (**1967**). ''Vergleichende Untersuchungen über Sexual- und Sozialverhalten der Krallenaffen (Hapalidae),'' Folia Primatogica **7**, 37–65.

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

Fitch, W. T. (**1994**). ''Vocal tract length perception and the evolution of language,'' Ph.D. thesis, Brown University.

Fitch, W. T., and Hauser, M. D. (**1995**). ''Vocal production in nonhuman primates: Acoustics, physiology and functional constraints on 'honest' advertising,'' Am. J. Primatology **37**, 191–219.

Fitch, W. T. (**in preparation**). ''Acoustic exaggeration of size in birds by tracheal elongation: comparative and theoretical analyses.''

Flügel, C., and Rohen, J. W. (**1991**). ''The craniofacial proportions and laryngeal position in monkeys and man of different ages. (A morphometric study based on CT-scans and radiographs),'' Mech. Aging Develop. **61**, 65–83.

Fox, M. W., and Cohen, J. A. (**1977**). ''Canid communication,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Fridlund, A. J. (**1994**). *Human Facial Expression: An Evolutionary View* (Academic, San Diego).

Gautier, J. P., and Gautier, A. (**1977**). ''Communication in Old World monkeys,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Geist, F. D. (**1933**). ''Nasal cavity, larynx, mouth, and pharynx,'' in *The Anatomy of the Rhesus Monkey*, edited by Hartman and Straus (Hafner, New York).

Goldstein, U. G. (**1980**). ''An articulatory model for the vocal tracts of growing children,'' D. Sc. thesis, Massachusetts Institute of Technology.

Goodall, J. (**1986**). *The Chimpanzees of Gombe: Patterns of Behavior* (Harvard U. P., Cambridge, MA).

Harrison, D. F. N. (**1995**). *The Anatomy and Physiology of the Mammalian Larynx* (Cambridge U. P., Cambridge, New York).

Harvey, P. H. (**1990**). ''Life-history variation: Size and mortality patterns,'' in *Primate Life History & Evolution*, edited by C. J. DeRousseau (Wiley–Liss, New York).

Hauser, M. D. (**1993**). ''The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight and social context,'' Am. Nat. **142,** 528–542.

Hauser, M. D., Evans, C. S., and Marler, P. (**1993**). ''The role of articulation in the production of rhesus monkey (*Macaca mulatta*) vocalizations,'' Animal Beh. **45,** 423–433.

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

Heinz, R. D., and Brady, J. V. (**1988**). ''The acquisition of vowel discriminations by nonhuman primates,'' J. Acoust. Soc. Am. **84,** 186–194.

Heinz, 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.

Hollien, H. (**1960**). ''Some laryngeal correlates of vocal pitch,'' J. Speech Hear. Res. **3,** 52–58.

Janis, C. (**1990**). ''Correlation and cranial and dental variables with body size in ungulates and macropodoids,'' in *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*, edited by J. Damuth and B. J. MacFadden (Cambridge U. P., Cambridge, England).

Joos, A. M. (**1948**). ''Acoustic phonetics,'' Language **24** (Suppl.), 1–136.

Kaufmann, J. H. (**1967**). ''Social relations of adult males in a free-ranging band of rhesus monkeys,'' in *Social Communication Among Primates*, edited by S. A. Altmann (University of Chicago, Chicago).

Kingdon, J. (**1974**). *East African Mammals: Vol. II Part A: Insectivores and Bats* (Academic, New York).

Künzel, H. J. (**1989**). ''How well does average fundamental frequency correlate with speaker height and weight?,'' Phonetica **46**, 117–125.

Ladefoged, P., and Broadbent, D. E. (**1957**). ''Information conveyed by vowels,'' J. Acoust. Soc. Am. **39,** 98–104.

Laitman, J. T., and Crelin, E. S. (**1976**). ''Postnatal development of the basicranium and vocal tract region in man,'' in *Symposium on Development of the Basicranium*, edited by J. F. Bosma (U. S. Government Printing Office, Washington, DC).

Lass, N. J., and Brown, W. S. (**1978**). ''Correlational study of speakers' heights, weights, body surface areas, and speaking fundamental frequencies,'' J. Acoust. Soc. Am. **63,** 1218–1220.

Lieberman, P. (**1968**). ''Primate vocalization and human linguistic ability,'' J. Acoust. Soc. Am. **44,** 1574–1584.

Lieberman, P. (**1975**). *On the Origins of Language* (Macmillan, New York).

Lieberman, P. (**1984**). *The Biology and Evolution of Language* (Harvard U. P., Cambridge, MA).

Lieberman, P., and Blumstein, S. E. (**1988**). *Speech Physiology, Speech Perception, and Acoustic Phonetics* (Cambridge U. P., New York).

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 **164,** 1185–1187.

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

Marler, P., and Tenaza, R. (**1977**). ''Signaling behavior of apes with special reference vocalization,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Martin, W. F. (**1972**). ''Evolution of vocalizations in the genus *Bufo*,'' in *Evolution in the genus Bufo*, edited by W. F. Blair (University of Texas, Austin, TX), pp. 279–309.

McComb, K. E. (**1991**). ''Female choice for high roaring rates in red deer, *Cervus elaphus*,'' Animal Beh. **41,** 79–88.

Miller, E. H. (**1975**). ''A comparative study of facial expressions of two species of pinnipeds,'' Behaviour **53**, 268–284.

Morita, S., and Ohtsuki, F. (**1973**). ''Secular changes of the main head dimensions in Japanese,'' Human Biol. **45**, 151–165.

Morton, E. S. (**1977**). ''On the occurrence and significance of motivation-structural rules in some bird and mammal sounds,'' Am. Nat. **111,** 855–869.

Nearey, T. (**1978**). *Phonetic Features for Vowels* (Indiana University Linguistics Club, Bloomington).

Negus, V. E. (**1949**). *The Comparative Anatomy and Physiology of the Larynx* (Hafner, New York).

Niemeier, M. M. (**1979**). ''Structural and functional aspects of vocal ontogeny in *Grus canadensis* (*Gruidae: Aves*),'' Ph.D. thesis, University of Nebraska, Lincoln.

Nowicki, S., and Marler, P. (**1988**). ''How do birds sing?,'' Music Percept. **5**, 391 426.

Ohala, J. J. (**1980**). ''ABSTRACT: The acoustic origin of the smile,'' J. Acoust. Soc. Am. Suppl. 1 **68,** S33.

Oppenheimer, J. R. (**1977**). ''Communication in New World monkeys,'' in

How Animals Communicate, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Owren, M. J. (**1990a**). ''Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: I. Natural calls,'' J. Comp. Psych. **104,** 20–28.

Owren, M. J. (**1990b**). ''Acoustic classification of alarm calls by vervet monkeys (*Ceropithecus aethiops*) and humans: II. Synthetic calls,'' J. Comp. Psych. **104,** 29–40.

Owren, M. J., and Bernacki, R. (**1988**). ''The acoustic features of vervet monkey (*Ceropithecus aethiops*) alarm calls,'' J. Acoust. Soc. Am. **83,** 1927–1935.

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

Parker, G. A. (**1974**). ''Assessment strategy and the evolution of fighting behavior,'' J. Theor. Biol. **47,** 223–243.

Peters, G. (**1984**). ''On the structure of friendly close range vocalizations in terrestrial carnivores (Mammalia: Carnivora: Fissipedia),'' Z. Säugetierkunde **49**, 157–182.

Peters, R. H. (**1983**). *The Ecological Implications of Body Size* (Cambridge U. P., New York).

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

Preuschoft, S. (**1992**). '' 'Laughter' and 'smiling' in Barbary macaques (*macaca sylvanus*),'' Ethology **91**, 220–236.

Preuschoft, S. (**1995**). '' 'Laughter' and 'smiling' in macaques: An evolutionary perspective,'' Ph.D. thesis, Utrecht University.

Pruit, C. H., and Burghardt, G. M. (**1977**). ''Communication in terrestrial carnivores: Mustelidae, Procyonidae and Ursidae,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Rendall, C. A. (**1996**). ''Social communication and vocal recognition in free-ranging rhesus monkeys (*Macaca mulatta*), Ph.D. thesis, University of California, Davis.

Richman, B. (**1976**). ''Some vocal distinctive features used by gelada monkeys,'' J. Acoust. Soc. Am. **60,** 718–724.

Ryalls, J. H., and Lieberman, P. (**1982**). ''Fundamental frequency and vowel perception,'' J. Acoust. Soc. Am. **72,** 1631–1634.

Ryan, M. J. (**1988**). ''Constraints and patterns in the evolution of anuran acoustic communication,'' in *The Evolution of the Amphibian Auditory System*, edited by B. Fritzch, M. J. Ryan, W. Wilczynski, T. E. Hetherington, and W. Walkowiak (Wiley, New York).

Schmidt-Nielsen, K. (**1984**). *Scaling: Why is animal size so important?* (Cambridge U. P., New York).

Schneider, R., Kuhn, H.-J., and Kelemen, G. (**1967**). ''Der Larynx des Männlichen *Hypsignathus monstrosus* Allen, 1861 (Pteropodidae, Megachiroptera, Mammalia),'' Z. wissenschaft. Zool. **175**, 1–53.

Schön, M. (**1971**). ''The anatomy of the resonating mechanism in howling monkeys,'' Folia Primatologica **15**, 117–132.

Senecail, B. (**1979**). *L'Os Hyoide: Introduction Anatomique a l'Etude de certains Mecanismes de la Phonation* (Faculté de Médicine de Paris, Paris).

Seyfarth, R. M., and Cheney, D. L. (**1984**). ''The acoustic features of vervet monkey grunts,'' J. Acoust. Soc. Am. **75,** 129–134.

Sharma, N. (**1990**). *Facial Growth and Body Size* (Agam Prakashan, Delhi).

Shipley, C., Carterette, E. C., and Buchwald, J. S. (**1991**). ''The effects of articulation on the acoustical structure of feline vocalizations,'' J. Acoust. Soc. Am. **89,** 902–909.

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.

Sullivan, B. K. (**1984**). ''Advertisement call variation and observations on breeding behavior of *Bufo debilis* and *B. punctatus*,'' J. Herpetology **18**, 406–411.

Tartter, V. C., and Braun, D. (**1994**). ''Hearing smiles and frowns in normal and whisper registers,'' J. Acoust. Soc. Am. **96,** 2101–2107.

Tuohimaa, P. T., Kallio, S., and Heinijoki, J. (**1981**). ''Androgen receptors in laryngeal cancer,'' Acta Oto-Laryngol. **91,** 159–164.

van Valkenburgh, B. (**1990**). ''Skeletal and dental predictors of body mass in carnivores,'' in *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*, edited by J. Damuth and B. J. MacFadden (Cambridge U. P., Cambridge, England).

van Dommelen, W. A. (**1993**). ''Speaker height and weight identification: a reevaluation of some old data,'' J. Phonetics **21**, 337–341.

van Hooff, J. A. R. A. M. (**1967**). ''The facial displays of the Catarrhine monkeys and apes,'' in *Primate Ethology*, edited by D. Morris (Weidenfeld and Nicolson, London).

van Hooff, J. A. R. A. M. (**1972**). ''A comparative approach to the phylogeny of laughter and smiling,'' in *Nonverbal Communication*, edited by R. Hinde (Cambridge U. P., New York).

Wakita, H. (**1976**). ''Instrumentation for the study of speech acoustics,'' in *Contemporary Issues in Experimental Phonetics*, edited by N. J. Lass (Academic, New York).

Walther, F. R. (**1977**). ''Artiodactyla,'' in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U. P., Bloomington, IN).

Walther, F. R. (**1984**). *Communication and Expression in Hoofed Mammals* (Indiana U.P., Bloomington, IN).