



Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags

DAVID REBY & KAREN MCCOMB

Experimental Psychology, School of Biological Sciences, University of Sussex

(Received 27 September 2001; initial acceptance 22 November 2001;
final acceptance 10 July 2002; MS. number: 7075R)

Early work on loud calling in mammals emphasized the importance of dynamic characteristics such as calling rate as cues to fitness and fighting ability. In contrast, little is known of the potential for fine-scaled acoustic cues to provide receivers with direct information on fitness. Fundamental frequency has typically been considered a good potential indicator of body size in the literature, but resonance frequencies (formants), which should be constrained by the length of the vocal tract, have received less attention. We conducted a detailed acoustic analysis on an extensive database of roars from red deer stags, *Cervus elaphus*, in a free-ranging population to investigate which variables provided honest information on age, body weight and reproductive success. Although fundamental frequency was higher in young stags than in adults, it did not decrease with body weight within adults and source cues (i.e. those generated by the larynx) in general did not provide clear information on fitness-related characteristics. In contrast, minimum formant frequencies, reached during the part of the roar when the mobile larynx is most fully retracted towards the sternum, decreased with body weight and age and were strongly negatively correlated with our index of reproductive success. Such production-related acoustic cues to body size and fitness, rendered honest by an anatomical constraint limiting the downward movement of the larynx, provide receivers with accurate information that could be used to assess rivals and choose mates.

© 2003 Published by Elsevier Science Ltd on behalf of The Association for the Study of Animal Behaviour.

Over the last 20 years, research on loud calling in mammals has focused on the importance of these vocalizations in determining the outcome of sexual competition, advertising territory ownership and coordinating the activities of members of a social group (e.g. Clutton-Brock & Albon 1979; Harrington & Mech 1979; Mitani 1987; East & Hofer 1991; McComb et al. 1994; Cowlshaw 1996; Reby et al. 1999a). However, perhaps because of the unavailability of appropriate tools for fine-scaled acoustic analysis, few studies have gone further than simply determining the context of call utterance and examining the effects of variation in caller identity and location, number of callers and basic dynamic features such as calling rate on the responses of receivers. The recent generalization of the 'source-filter' theory of voice production (Fant 1960) to mammals other than humans (primates: Lieberman et al. 1969; Owren 1990; Fitch 1997; cats, *Felis catus*: Carterette et al. 1979; dogs, *Canis familiaris*: Riede & Fitch 1999; reviewed in Fitch & Hauser, in press) has generated the opportunity to study how, in the light of their biomechanical modes of production, acoustical

features of vocal signals can provide receivers with direct information on physical characteristics of the caller.

The key element of source-filter theory is that any voiced signal results from a source signal (generated by vibration of the vocal folds in the larynx) being subsequently filtered in the cavities of the vocal tract (Fant 1960). The source signal, typically a quasiperiodic wave with a fundamental frequency and integer multiple harmonics, initially determines the pitch of a voiced vocalization. This signal then passes through the supralaryngeal vocal tract which, as a tube of air with natural resonances, selectively amplifies certain frequencies and finally radiates out through the mouth and nostrils into the environment. This filtering process thus shapes the spectral envelope of the signal, producing peaks called 'formants' (Fant 1960). Since characteristics of vocalizations that arise from inherent properties of the filter can vary independently from those that arise from the source, either or both may provide receivers with important information on the caller's physical stature and condition (Fitch & Hauser, in press).

Until recently, where researchers working on mammals have considered how information on the callers' attributes are coded in vocalizations, they have tended to

Correspondence: D. Reby, Experimental Psychology, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, U.K. (email: reby@biols.susx.ac.uk).

ignore the effects of the filter and focus instead on the potential importance of source-related pitch (e.g. Masataka 1994). The most influential work in this area has been based on the prediction that fundamental frequency should provide an accurate cue to body size (Morton 1977), with larger animals producing lower-pitched vocalizations. This prediction rests on the assumptions that the length of the vocal folds increases with body size and constrains fundamental frequency range (Titze 1994). The prediction that fundamental frequency should decrease with body size across species has been verified in nonhuman primates, with larger species producing relatively lower-pitched vocalizations than smaller species (Hauser 1993). However, the prediction that within a species, larger individuals should produce vocalizations with a lower fundamental frequency than smaller individuals has not stood up well to empirical test in loud-calling mammals (McComb 1991; Masataka 1994).

A number of studies that have examined what information may be conveyed to receivers in the filter-related formants suggest that this feature may provide a more reliable indication of body size (red deer, *Cervus elaphus*: McComb 1988; rhesus monkeys, *Macaca mulatta*: Fitch 1997; dogs: Riede & Fitch 1999). This is a direct consequence of the physical relation that links formant frequencies to the length of the vocal tract modelled as a quarter wave resonator (Fant 1960; Titze 1994). Schematically, formant frequencies and overall formant dispersion should decrease when the length of the vocal tract increases (Fitch 1997). Since in most mammals the larynx is tightly attached to the base of the skull, the length of the vocal tract is constrained by skull size (Fitch 1997, 2000; Fitch & Hauser, *in press*). Therefore, formant frequencies have the potential to reflect variation in body size accurately.

Red deer males have a descended larynx (Fitch & Reby 2001), and are able to drop their highly mobile larynx further down towards the sternum, which allows them to increase their vocal tract length while they vocalize. Red deer stags roar repeatedly during the autumn breeding season (rut), and roaring is known to affect the outcome of male–male contests, influence mate attraction and advance ovulation (Clutton-Brock & Albon 1979; McComb 1987, 1988, 1991). Fitch & Reby (2001) have hypothesized that stags extend their vocal tract to decrease formant frequencies and spacing and thus give receivers an inflated impression of their actual body size, increasing the chances that they will attract mates and deter competitors. This anatomical adaptation suggests that there has been strong selection over evolutionary time for broadcasting size-related information in this way. However, it also raises the possibility that formant dispersion may no longer be a reliable, ‘honest’ indicator of body size, since individual stags can actively modify their vocal tract length independently from their body size. It is important to test whether pulling the larynx down as far as the sternum (Fitch & Reby 2001) imposes an upper limit on vocal tract length, which could constrain the lowest frequency of the formants and consequently their minimum spacing in the roar. If so, the

lowest frequency of each individual formant and the minimum frequency spacing of the formants, by reflecting the maximum extension of the vocal tract during roaring, may still provide listeners with honest information.

We investigated which source- and filter-related acoustic features in red deer roars have the potential to convey information on physical attributes of callers. To do this, we analysed roars of 57 red deer stags of known age (body weight and/or reproductive success were also known for 24 of these stags) to investigate the variation of selected spectral characteristics with age and body size. We examined how the range and stability of the fundamental frequency contour vary with age, body weight and reproductive success. In addition, we investigated the relation between age, body weight, reproductive success and the position of the formants at maximum vocal tract extension.

METHODS

Recording Red Deer Roars

Roars were recorded from red deer in a study population on the island of Rum (Inner Hebrides, U.K.), where stags have been intensively studied for 30 years (Clutton-Brock et al. 1982; Kruuk et al. 2000). Recordings were made with Sennheiser MKH 816 and MKH 416 microphones linked to a Uher 4200 Report Monitor open reel, a Marantz CP 230 cassette recorder or a HHB PDR 1000 professional DAT recorder. We considered roars recorded between 1976 and 1999, from a set of 57 stags of known age (age range 5–13 years, mean 8.3) subdivided into two age subsets: subadult (5–6 years, $N=10$) and adult (7–13 years, $N=47$). The adult subset included 24 harem-holding adult stags (age range 7–10 years, mean 8.7) whose weight or reproductive success was known, recorded between 1981 and 1986 (McComb 1987, 1988, 1991).

Prerut Body Weight as Body Size Index

Twenty-two of the stags described above had been darted and weighed in the field when they were at least 5 years old. All dartings were conducted with an Injekta darting system (Norinject, Holen, Norway), which uses lightweight plastic darts propelled by compressed air. The darts were fired from up to 50 m away, into the haunch muscle, and the drug (Large Animal Immobilon, containing Etorphine, 0.008–0.01 ml/kg) was automatically injected on impact. The immobilization time was always less than 40 min. The weighing equipment consisted of a clock-style balance attached to an aluminium tripod. The deer was laid on a canvas triangle that was attached to the balance hook. After the injection of the reversal agent (Large Animal Revivon), the recovery time was 3–4 min.

Because darting was not conducted in the year and month of the recording, and because body weight varies with age and season, we estimated the prerut (August) body weight for each of these stags for the year in which

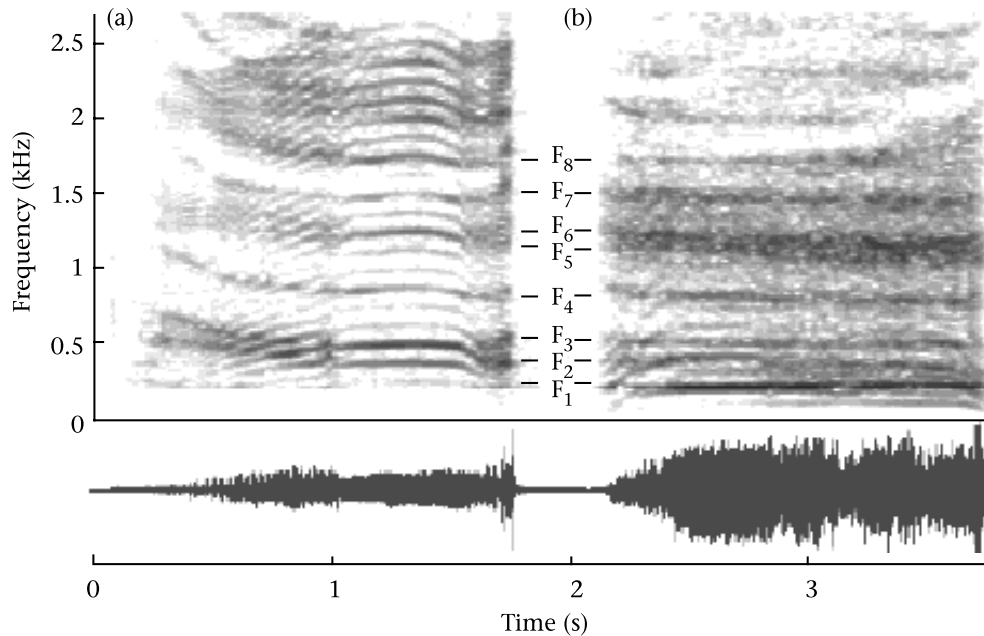


Figure 1. Waveforms and spectrograms of common and harsh roars from a red deer stag. (a) Common roars are harmonically rich with modulated source (fundamental frequency) and filter (formants). The narrow, evenly spaced frequency bands rising in the first half of the roar and decreasing towards the end represent the fundamental frequency and harmonics. The vocal tract resonances, or 'formants' are represented by the darker, less regularly spaced energy bands (labelled as F_1, F_2, \dots, F_8) that decrease during calling, as the stag lengthens its vocal tract. Harmonics and formants move independently because of source-filter independence. (b) Harsh roars are generally louder, and characterized by poorly defined or no harmonics and little or no formant modulation. In both common and harsh roars, spacing between successive formants is not perfectly even, suggesting that the red deer supralaryngeal vocal tract departs slightly from the straight tube model.

its roars were recorded. This estimate was based on weight increments per year and weight deviations across months predicted from growth curves constructed from actual age-specific weights obtained from stags present on Rum during the study period. The range of estimated weights was 89–150 kg (mean 125 kg). Since red deer stags lose weight during the rut (Mitchell et al. 1976), prerut estimates of body weights cannot be regarded as indicating exact body weights on the days roars were recorded. However, they constitute good indirect indicators of differences between individuals in body size, which is the underlying factor that should affect the structure of roars.

Reproductive Success Indices

We obtained a relative index of reproductive success for 22 adult stags (20 of those included in the weight sample and two additional stags). Reproductive success in any year was estimated by the number of 'hind-days held' in the 10 days on either side of the median conception date for that year (21-day window), where 1 hind-day represents tenure of one hind for 1 day (Clutton-Brock et al. 1982). We repeated this for the years in which the stag was aged 7, 8, 9 and 10 years (spanning the most successful years of a stag's reproductive life). The number of hind-days held was normalized by log transformation and entered as the dependent variable in an analysis of variance with identity and age as factors. Multiple classification then expressed log hind-days held for each of the stags as a deviation from the grand mean of the variable across all 22 stags in the analysis. We used these devi-

ations as an age-independent index of reproductive success. Reproductive success indices calculated on this basis were strongly positively correlated with estimated body weights ($r^2=0.45, N=20, P=0.0012$).

Description of Call Types

The red deer roar is generally composed of three distinct phases, reflecting changes in vocal fold vibration and vocal tract shape occurring during sound production (McComb 1988; Reby 1998; Fitch & Reby 2001). In the first phase of the roar, the fundamental frequency typically rises (Fig. 1a), reflecting an increase in the rate of vibration of the vocal folds probably caused by the increasing subglottal pressure and vocal fold tension. Simultaneously, the stag increases the length of its vocal tract by lowering its larynx and raising its head, thus decreasing formant frequencies (Fitch & Reby 2001). During the second phase, the vocal tract of the animal usually remains extended and as a consequence formant frequencies reach their minimum values (as shown in Fitch & Reby 2001). In contrast, the amplitude and the fundamental frequency are relatively high during this part of the call. This phase is also sometimes characterized by segments with broadband chaotic noise, in which little or no harmonic structure is visible, indicating non-periodic, chaotic vibration of the vocal folds. During the last phase of the roar, which is typically shorter, amplitude and fundamental frequency fall rapidly while formants rise. These changes arise from decreases in subglottal pressure and vocal fold vibration rate, and the

larynx and head coming back to their initial resting positions.

A minority of roaring bouts include vocalizations known as 'harsh roars' (Fig. 1b). Narrow band spectrograms show that the acoustic structure of harsh roars is similar to that of the noisiest segments of the more typical roars described above (which we refer to as 'common roars') with the fundamental frequency and harmonics poorly defined or absent. Harsh roars are also characterized by weaker formant modulation, the absence of a pronounced drop in formant frequencies at the beginning of the roar, reflecting the static body posture adopted by the animal during the production of this call. In harsh roars the stag fully extends its neck (the head is raised so that the lower jaw is aligned with the sternal limit). Comparison of the intensity of common and harsh roars uttered in the same bout (with constant recording conditions) suggests that the harsh roars are generally louder than common roars. Harsh roars tend to occur in situations of intense activity, after a roaring contest or during a period of repeated herding. Bouts delivered in these situations often start with a series of short roars or 'grunts' followed by one or more harsh roars.

Selection of Roar Bouts for Analysis

For each of the 57 stags in the study, we edited 5–10 (average 9.3) bouts of common roars given during the peak of vocal activity, when stags were actively holding harems and challenged by surrounding stags. Preliminary observations suggested that stags put more effort into the production of the first roar of each bout. The first roar is generally the loudest and highest-pitched (i.e. with highest fundamental frequency) call in the bout. During this roar, stags lower their larynx towards the sternal limit and maintain their vocal tract extension over long portions of the roar. As a result, each formant in this call has a well-defined minimum plateau that reflects maximum vocal tract length achieved during roaring. Given the highly variable number of roars per bout within individuals (1–11), and the fact that the plateau reflecting maximal vocal tract extension is not always reached in later roars, we standardized our comparisons by considering only the first roar of each bout in our acoustical analyses. Although it is possible that frequency components in the subsequent roars may encode additional information on physical characteristics of the caller (see Discussion), we do not present analyses on these roars here.

We also examined the formant frequencies of one bout of 'harsh roars' in nine adult stags for which we had estimated body weight and reproductive success indices. In these cases the roar in which formant frequencies reached their lowest position was selected for analysis.

Vocalizations were digitized with an Audiomedia III sound card connected to a power PC 7600 Macintosh computer. Each call was imported to the computer with Sound-Edit 16V2 and saved in AIFF format, at 16 bits amplitude resolution and 22 kHz sampling rate. DAT

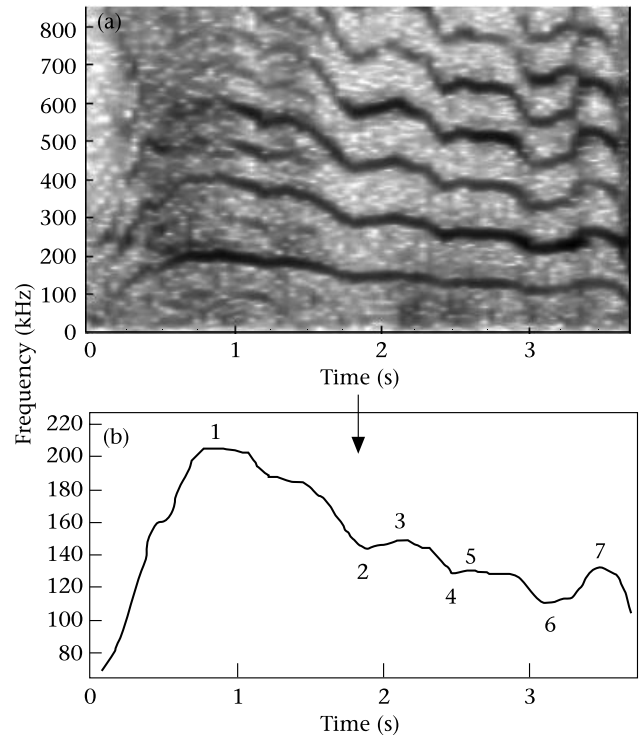


Figure 2. Extraction of source-related acoustic variables in a red deer roar. (a) Spectrogram of a roar showing the fundamental frequency and the first harmonics. (b) Pitch extraction for this roar. The fundamental frequency range is 69–206 Hz with an average of 147 Hz. There are seven inflexion points (numbered 1–7) for a duration of 3.75 s (1.9 inflexions/s) and a cumulated fundamental frequency variation of 319 Hz (86.7 Hz/s).

recordings were directly transferred to the computer hard disk using the card SPDIF port. Sound files were down-sampled to 11 kHz after appropriate low-pass filtering, and narrow-band spectrograms (FFT size=512, overlap=50%, frequency grid resolution=21.53 Hz) of each call were edited and saved with Canary 1.2.4 software (Chariff et al. 1995).

Acoustic Analyses

We extracted source- (fundamental frequency) and filter-(formant) related acoustic features with the Praat 3.9.27 DSP package (P. Boersma & D. Weenink, University of Amsterdam, The Netherlands).

To characterize the source (Fig. 2), we measured several features from the fundamental frequency contour. A forward autocorrelation (To pitch (cc) command) algorithm was used to produce time-varying numerical representations of the fundamental frequency, F_0 . The time step in the analysis was 0.04 s. With this algorithm, it is necessary to specify for each call the frequency range within which F_0 is expected to vary, to increase the accuracy of the output results by limiting the number of 'octave jumps'. The limits of this range (minimum and maximum F_0) were determined by visual assessment of the variation of the fundamental frequency in the call, using narrow bandwidth spectrographic analysis in Praat

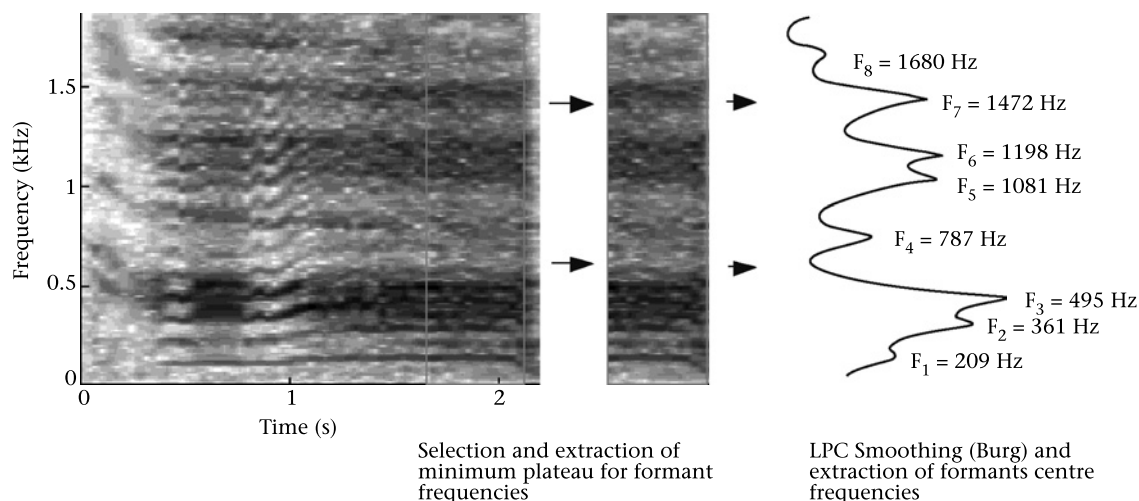


Figure 3. Extraction of the minimum frequency of individual formants in a red deer roar. The section of the spectrogram where individual formants reach their lowest frequencies in the roar was selected and the corresponding sound wave was saved into a separate file for LPC (Linear Predictive Coding) analysis. The centre frequencies of the formants were averaged on the duration of this section.

(window length 0.05 s). The contour of the fundamental frequency was inferred from on-screen spectrographic examination of the second or third harmonics (thereby excluding any underestimation of the minimum values of F0 caused by a possible attenuation of the frequencies close to the roll-off frequency range of our recording equipment, or by background noise). Typical preset values for the lower and upper limits of the expected range for F0 were 30 and 250 Hz, respectively. The time window in the frequency analysis was variable and automatically imposed by the preset lower limit for fundamental frequency.

We wrote a script in Praat to batch process the editing, the setting of analysis parameters, the analyses and the exporting of output data. The output numerical representation of the frequency contour was automatically saved in text format and transferred to Microsoft Excel to derive the following measurements with automated macros.

(1) Fundamental frequency: minimum, mean and maximum in the call (minF0, meanF0, and maxF0).

(2) The stability and amplitude of variation of the fundamental frequency contour were assessed with the following indices: (a) $\text{inflex} = i/T$: the average number of inflexions in fundamental frequency contour per unit time (Fig. 2), where i is the number of inflexions of the fundamental frequency contour over the duration of the call and T is the duration of the call. The parameter i was calculated as the number of changes in the sign of the derivative of the fundamental frequency contour, after running a five-point average smoothing filter to remove rapid variations caused by jitters or analysis imprecision. (b) $\text{sumvar} = \Delta F0/T$: the average fundamental frequency variation per unit time (Fig. 2), where $\Delta F0$ is the cumulated variation of F0 over the duration of the call and T is the duration of the call. $\Delta F0$ was estimated as the sum of the absolute value of the F0 derivative.

To characterize the filter, we measured the minimum frequency values of the first eight formants (minF1,

minF2 . . . minF8) using Linear Predictive Coding (LPC: 'To Formants (Burg)' command in Praat, Press et al. 1992). As described earlier, formant frequencies in red deer roars decrease as a stag lowers its larynx to extend its vocal tract, until they reach a minimum plateau corresponding to the maximum vocal tract extension, as constrained by the larynx approaching the level of the sternum. We assessed the temporal limits of this plateau from spectrograms and selected and saved the corresponding sound wave section (Fig. 3). Formant analysis parameters were set as follows: time step: 0.04 s; maximum number of formants: 8–10; maximum formants: 2000–2300 Hz; window analysis: 0.1 s. Formant frequency values were transferred to Excel and averaged over the duration of the plateau.

We then estimated the maximum vocal tract length during roaring (maxVTL) from the minimum frequencies of the formants of each stag. The vocal tract was approximated as a straight uniform tube closed at one end (the glottis) and open at the other end (the mouth). The centre frequencies of the successive formants (F_1, F_2, \dots, F_i) generated by such a resonator are related to the length of the vocal tract by the equation:

$$F_i = \frac{(2i-1)c}{4\text{VTL}} \quad (1)$$

where c is the speed of sound in air (approximated as 350 m/s in a mammal's vocal tract) and VTL is the length of the vocal tract (Titze 1994).

The spacing between any two consecutive formants in the frequency spectrum is therefore constant and given by:

$$\Delta F = F_{i+1} - F_i = \frac{c}{2\text{VTL}} \quad (2)$$

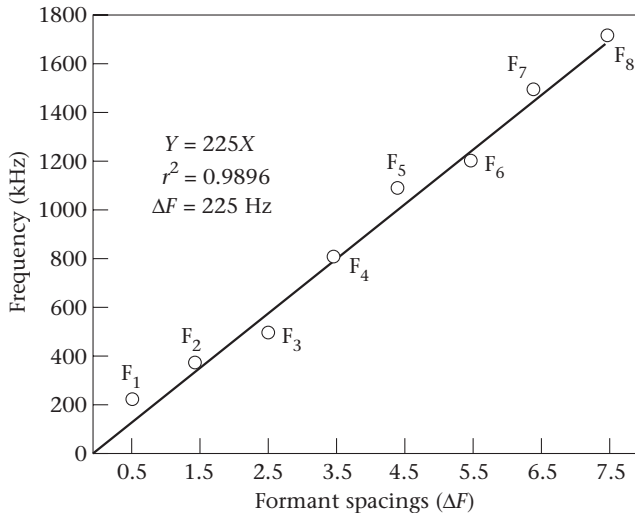


Figure 4. Illustration of the method used to estimate minimum spacing of formants during roaring for a particular stag. Observed minimum frequency values of each formant (minF_1 , to minF_i) were plotted against $(2i-1)/2$ increments of the formant spacing as predicted by our vocal tract model. Then, a linear regression line was fitted to the set of observed values, using an intercept equal to 0. Since $F_i = 2i-1/2\Delta F$, the slope of the linear regression (in bold) gives the best estimate of $\text{min}\Delta F$ for our vocal tract model. The estimated maximum vocal tract length of this stag during roaring (maxVTL) was finally deduced from the estimate of $\text{min}\Delta F$ by applying equation (4) (see text). In this case, $\text{min}\Delta F = 225$ Hz, which corresponds to an estimated maximum vocal tract length of 77.8 cm.

Furthermore, by replacing

$$\frac{c}{2\text{VTL}}$$

by ΔF in equation (1), individual formant frequencies can be related to formant spacing ΔF by the equation:

$$F_i = \frac{2i-1}{2} \Delta F \quad (3)$$

For each stag, we estimated $\text{min}\Delta F$, the minimum spacing of the formants, by seeking the best fit for equation (3) to the lowest frequency of each formant measured in this stag's roars (minF_1 to minF_8 ; Fig. 4).

An estimate of the maximum vocal tract length of the stag during roaring, maxVTL , was then simply deduced from $\text{min}\Delta F$ by:

$$\text{maxVTL} = \frac{c}{2(\text{min}\Delta F)} \quad (4)$$

Statistical Analyses

The effect of age on each acoustic parameter was studied on the data set of 57 stags, using simple linear regression. In addition, t tests were computed on two subsets: subadult stags (aged 5–6 years, $N=10$) and adults ($N=47$) to test for differences between these two age classes. The effects of weight and reproductive success

were studied for adults for which we had estimated weights and/or indices of reproductive success ($N=24$), using simple linear regressions. We also conducted a multiple regression to partial out the effects of age and weight on acoustic parameters within the adult data set. Since weight is linearly related to volume, and since volume is proportional to the cube of length, body weight estimations were log transformed to convert the cubic relation between length-related variables (minF_i , $\text{min}\Delta F$ and maxVTL) and weight into a linear one.

RESULTS

Voice Characteristics Versus Age

Fundamental frequency variables

Roar pitch was a poor predictor of age in the overall data set representing stags of 5–13 years old. Among the source-related variables we measured, minimum, mean and maximum fundamental frequency were not significantly correlated with age (Table 1), although mean and maximum fundamental frequencies were significantly higher in subadults than in adults (Table 1). Nor was there a relation between age and the number of inflexion points per unit time (inflex), or the cumulated fundamental frequency modulation per unit time (sumvar ; Table 1), and neither of these variables differed between adults and subadults (Table 1).

Formant variables

In contrast, formant frequencies provided a clear indication of the age of the caller. All of the minimum formant frequencies we measured (minF_1 , . . . minF_8) as well as the minimum frequency spacing of formants ($\text{min}\Delta F$) were negatively correlated with age (Table 1, Fig. 5a). These filter-related variables were also significantly higher in subadults than in adults (Table 1). The range of estimated maximum vocal tract lengths during roaring (maxVTL) was 60.5–78.0 cm (average 70.7); length increased with age (Fig. 6a), and on average vocal tracts were 6.8 cm longer in adults than in subadults.

Voice Characteristics Among Adults

Fundamental frequency variables

Roar pitch did not provide information on body weight amongst adult stags. None of the fundamental frequency values we measured (minF_0 , meanF_0 or maxF_0), nor the fundamental frequency stability indices (inflex and sumvar), was significantly correlated with body weight (Table 2).

Formant variables

In contrast, formant frequencies provided good information on body weight among adult stags (Table 2). minF_1 , minF_6 , minF_7 , minF_8 and $\text{min}\Delta F$ were all negatively correlated with body weight (Table 2, Fig. 5b) and maxVTL was positively correlated with body weight (Fig. 6b). Removing from the analysis the outlying stag (89 kg) increased the slope of the regression line (Fig. 6b).

Table 1. Regression of roar acoustic variables with age across life span in red deer stags, and comparison of means between subadults and adults

Variable	Regression with age (N=57)		Comparison of means			
	<i>r</i>	<i>P</i>	Subadults (N=10)	Adults (N=47)	<i>t</i>	<i>P</i>
Fundamental frequency						
MinF0	-0.059	0.665	70.8	61.7	1.83	0.073
MeanF0	-0.192	0.152	125.1	106.9	2.08	0.042
MaxF0	-0.232	0.083	162.4	136.8	2.14	0.037
Inflex	0.170	0.207	1.8	1.9	-0.93	0.355
Sumvar	-0.028	0.834	77.9	77.2	0.08	0.937
Formant frequencies						
MinF1	-0.679	0.0001	241.4	208.2	6.03	0.0001
MinF2	-0.308	0.020	446.1	399.0	3.18	0.002
MinF3	-0.525	0.0001	634.4	556.5	6.02	0.0001
MinF4	-0.706	0.0001	992.0	874.9	7.44	0.0001
MinF5	-0.589	0.0001	1261.3	1178.1	5.90	0.0001
MinF6	-0.752	0.0001	1452.2	1305.9	10.34	0.0001
MinF7	-0.725	0.0001	1737.2	1574.6	7.72	0.0001
MinF8	-0.699	0.0001	2000.2	1802.9	7.25	0.0001
MinΔF	-0.755	0.0001	269.2	243.5	9.04	0.0001
MaxVTL	0.762	0.0001	65.1	71.9	-8.69	0.0001

Inflex: number of inflexion points per unit time; sumvar: cumulated variation per unit time; minΔF: minimum formant spacing; maxVTL: maximum estimated vocal tract length.

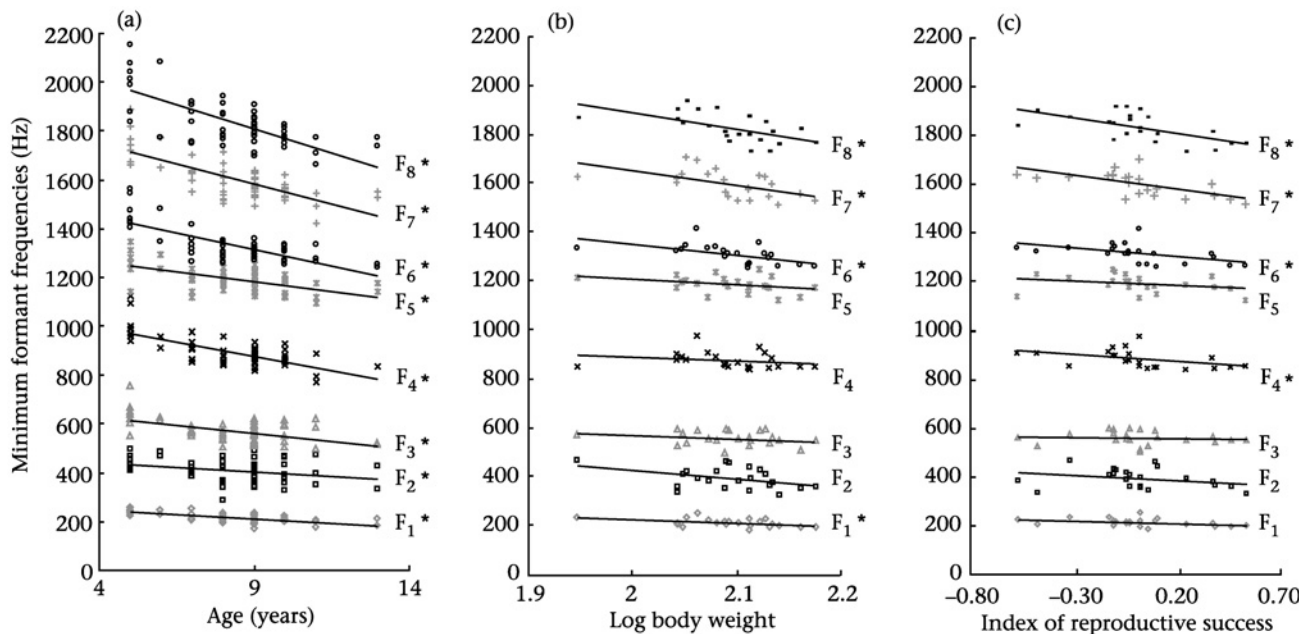


Figure 5. Simple regressions of the minimum frequency of each formant in red deer roars with (a) age across life span and, for adults, (b) body weight and (c) reproductive success. * $P < 0.05$. See Tables 1, 2 and 4 for details of the regression statistics.

Although we used body weight, body size would theoretically constitute a better predictor of vocal tract length. The indirect nature of the relation between body weight and size-related acoustic variables, and inaccuracies inherent to our prerut weight estimation procedure, probably explain some of the remaining variation in formant frequencies (which is independent of body weight).

In our sample of harsh roars, minF4, minF5, minF6 and minΔF were all negatively correlated with body weight (Table 3). In addition, maximum vocal tract length during harsh roaring increased with weight (Table 3). In the nine stags for which we had recordings of both common and harsh roars, estimated maximum vocal tract length was longer during harsh roars than during common roars (paired t test: $t_{17} = 120.5$, $P < 0.0001$), corresponding to an

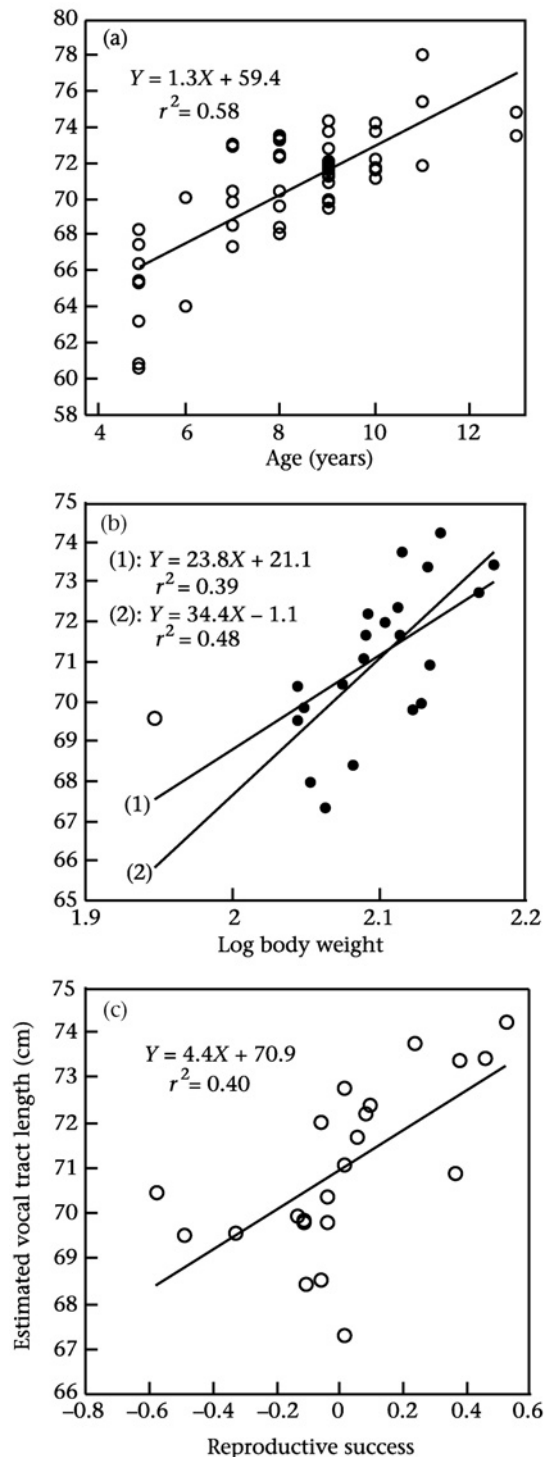


Figure 6. Simple regressions of maximum vocal tract length (maxVTL) with (a) age across life span and, for adults, (b) estimated body weight and (c) reproductive success. Estimated maximum vocal tract length increased significantly with age ($r=0.762$, $N=57$, $P=0.0001$), body weight ($r=0.625$, $N=22$, $P=0.0019$) and reproductive success ($r=0.647$, $N=22$, $P=0.0011$). In (b), the solid line marked (2) was obtained by removing the stag with an abnormally light body weight (○) from the analysis. This removal increased the slope and significance of the regression ($r=0.69$, $N=21$, $P=0.0005$).

average increase of 1.8 cm in the estimated maximum vocal tract length.

Effects of age and weight on roar variables

We used multiple regression to partial out the effects of age and body weight on roar variables within the adult data set of common roars. Within adults, controlling for the effects of age did not alter the relation between any of the acoustic variables and body weight; as before, formants were good predictors of body weight in contrast to fundamental frequency variables, which were not (Table 2). However, when weight was controlled for among adults, minF0 and maxF0 were negatively correlated with age (Table 2). In addition, min Δ F appeared to decrease with age and maxVTL to increase ($P=0.06$) with age as in the life span data set (Table 2).

When the effects of age and body weight were partialled out using multiple regression in our smaller data set of harsh roars, the results were the same as in the simple regression (see above and Table 3).

Voice Characteristics and Reproductive Success

Fundamental frequency variables

MinF0 showed a positive correlation with our index of reproductive success (rather than negative as predicted), but not meanF0 or maxF0 (Table 4). Neither inflex nor sumvar was correlated with reproductive success (Table 4).

Formant variables

MinF4, minF6, minF7, minF8 and min Δ F estimated from common roars all decreased with increasing reproductive success (Table 4, Fig. 5c), while maxVTL increased (Fig. 6c). In our smaller sample of harsh roars from stags whose reproductive success was known, the minimum frequency of each individual formant, min Δ F and maxVTL were not correlated with our index of reproductive success (Table 4).

DISCUSSION

Our results indicate that in adult red deer stags fundamental frequency does not provide a clear indication of body size, despite the existence of a positive anatomical correlation between vocal fold length, age and body weight (W. T. Fitch, D. Reby, B. Cargnelutti & K. McComb, unpublished data). Fundamental frequency in the first roar of the bout was higher on average in subadult stags, probably reflecting an increase in vocal fold length in stags across their lifetime and there was a negative correlation between some fundamental frequency variables and age among adults when the effects of weight were controlled for. However, we found a strong variability in F0 among adults, including values typical of young animals or even higher. In adults, this variability was independent of body weight (and therefore presumably of vocal fold length) and within adults minimum fundamental frequency was positively (rather than negatively) correlated with reproductive success.

Table 2. Descriptive statistics for acoustic variables in the roars of adult red deer stags, simple regressions with body weight and multiple regressions with age and body weight

Variable	Descriptive statistics (N=22)			Single regression with log weight (N=22)		Multiple regression			
	Minimum	Mean	Maximum	<i>r</i>	<i>P</i>	Partial regression with age (N=22)		Partial regression with log weight (N=22)	
						<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
Fundamental frequency									
MinF0 (Hz)	36.3	64.6	92.6	0.248	0.266	0.119	0.609	0.287	0.225
MeanF0 (Hz)	65.7	111.7	168.3	0.086	0.705	-0.446	0.052	0.211	0.338
MaxF0 (Hz)	71.1	142.2	213.9	0.030	0.893	-0.480	0.035	0.163	0.450
Inflex	0.9	1.9	2.8	0.345	0.116	0.315	0.154	0.257	0.240
Sumvar (Hz/s)	18.6	72.1	126.7	0.023	0.921	-0.432	0.061	0.143	0.519
Formant frequencies									
MinF1 (Hz)	186	213	252	-0.484	0.022	0.229	0.252	-0.498	0.019
MinF2 (Hz)	329	396	469	-0.400	0.065	-0.119	0.576	-0.393	0.075
MinF3 (Hz)	502	559	600	-0.227	0.311	-0.100	0.658	-0.221	0.334
MinF4 (Hz)	842	881	980	-0.224	0.317	-0.050	0.826	-0.221	0.336
MinF5 (Hz)	1124	1190	1253	-0.358	0.102	-0.135	0.534	-0.350	0.116
MinF6 (Hz)	1260	1314	1416	-0.572	0.005	-0.107	0.574	-0.566	0.007
MinF7 (Hz)	1516	1602	1714	-0.556	0.007	-0.052	0.789	-0.553	0.009
MinF8 (Hz)	1733	1830	1940	-0.580	0.005	-0.151	0.424	-0.571	0.006
MinΔF (Hz)	236	247	260	-0.621	0.002	-0.342	0.059	-0.526	0.006
MaxVTL (cm)	67.3	70.6	74.2	0.625	0.002	0.334	0.064	0.532	0.006

Inflex: number of inflexion points per unit time; sumvar: cumulated variation per unit time; minΔF: minimum formant spacing; maxVTL: maximum estimated vocal tract length.

Table 3. Harsh roar filter variables in adult stags: simple regressions with body weight and multiple regressions with age and body weight

Variable	Simple regression with log weight (N=9)		Multiple regression			
	<i>r</i>	<i>p</i>	Partial regression with age (N=9)		Partial regression with log weight (N=9)	
			<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
MinF1	-0.375	0.320	-0.305	0.431	-0.335	0.389
MinF2	-0.268	0.486	0.166	0.688	-0.275	0.511
MinF3	-0.327	0.391	-0.125	0.757	-0.315	0.446
MinF4	-0.842	0.004	-0.286	0.158	-0.817	0.004
MinF5	-0.743	0.022	-0.309	0.255	-0.699	0.030
MinF6	-0.727	0.027	0.252	0.373	-0.767	0.026
MinF7	-0.443	0.232	-0.258	0.492	-0.416	0.283
MinF8	-0.600	0.088	-0.339	0.298	-0.557	0.111
MinΔF	-0.771	0.015	-0.268	0.296	-0.740	0.020
MaxVTL	0.768	0.016	0.265	0.307	0.738	0.021

These roars are characterized by a nonperiodical source and so no fundamental frequency variables were extracted. MinΔF: minimum formant spacing; MaxVTL: maximum estimated vocal tract length.

This suggests that other variables, including factors associated with physical fitness, may be involved in counteracting the expected effect of vocal fold length on roar fundamental frequency.

When mammalian vocal folds are modelled as strings, their fundamental frequency can be predicted on the basis of the length of their vibrating portion, tissue density and longitudinal stress (Titze 1994). In red deer stags it cannot be assumed that tissue density and longitudinal stress would be constant across individuals.

Furthermore, fundamental frequency is also influenced by subglottal pressure. High-amplitude oscillation of the vocal folds, generated by increasing subglottal pressure, increases the vocal fold strain ('dynamic strain', Titze 1994, pp. 209–210) and leads to an increase in fundamental frequency. It is possible that in red deer, stronger stags (with stronger chest muscles and larger pulmonary capacity) or more highly motivated stags (roaring at greater volumes) could produce higher subglottal pressures. High motivation, characteristic of stags with larger

Table 4. Simple regressions of roar acoustic variables with reproductive success in adult red deer stags for both common and harsh roars

Variable	Common roars (N=22)		Harsh roars (N=8)	
	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
Fundamental frequency				
MinF0	0.464	0.013		
MeanF0	0.265	0.233		
MaxF0	0.210	0.347		
Inflex	0.342	0.119		
Sumvar	0.122	0.588		
Formant frequencies				
MinF1	-0.406	0.059	-0.617	0.103
MinF2	-0.283	0.201	-0.272	0.514
MinF3	-0.112	0.621	-0.124	0.769
MinF4	-0.437	0.042	-0.640	0.088
MinF5	-0.274	0.217	-0.360	0.382
MinF6	-0.510	0.015	-0.621	0.100
MinF7	-0.633	0.002	-0.311	0.454
MinF8	-0.613	0.002	-0.396	0.332
MinΔF	-0.629	0.002	-0.532	0.175
MaxVTL	0.647	0.001	0.528	0.179

Inflex: number of inflexion points per unit time; sumvar: cumulated variation per unit time; minΔF: minimum formant spacing; maxVTL: maximum estimated vocal tract length.

harems, might also result in higher fundamental frequencies by other mechanisms, for example by neural control resulting in greater vocal fold stiffness. We suggest that the high variability in fundamental frequency observed in adult stags results from conflicting factors: F0 may decrease with vocal cord length, but increase with factors such as subglottal pressure, which is likely to be higher in stronger animals with larger body sizes and higher reproductive success. This would also explain the positive correlation we observed between minimum fundamental frequency and stag reproductive success and the general tendency for all fundamental frequency values to increase with increasing stag body weight and reproductive success.

In the Cervid family, there is an apparent lack of correlation between species (or subspecies) body size and fundamental frequency. Indeed, whereas both fallow deer, *Dama dama*, and Corsican deer, *C. e. corsicanus*, are smaller than red deer, the fundamental frequency of the rutting calls of the males is much lower in both species than in red deer (fallow deer: 35 Hz, Reby et al. 1998; Corsican deer: 34 Hz, B. Cargnelutti & D. Reby, unpublished data). Similarly, in the American subspecies of red deer, the wapiti, *C. e. canadensis*, the typical fundamental frequency of bulls' rutting calls (bellows) is around 1 kHz, although on average they weigh 2.5 times as much as red deer. Although these apparently counterintuitive observations may reflect a specialization of the shape and histology of the vocal folds, or of a lack of correlation between body size and larynx size across species, they emphasize that the multiple factors determining fundamental frequency may interact in too complex a way to

provide clear information about size across and within species.

Other characteristics of the source, describing the stability of the fundamental frequency, could theoretically reflect the ability of the sender to maintain adequate neuromuscular control of respiration and phonation (Titze 1994; Hauser 1996). However, we found that the cumulated variation and number of inflexion points of the fundamental frequency contour per unit time were not correlated with any of the factors considered, suggesting that no information regarding the quality of the caller is advertised in these variables.

In contrast to the limited information available in the fundamental frequency variables discussed above, formant frequencies provided reliable cues to stag age, body size and reproductive success. Formant frequencies and spacing decreased with age, probably because of the lengthening of the neck and vocal tract that occurs with body growth. This decrease may also result from an increase in the elasticity of the thyrohyoid ligaments, allowing adults to be more efficient than subadults at fully extending their vocal tract by lowering the larynx closer to the sternum limit. Within adults, we found that half of the formant frequencies and also formant spacing were strongly correlated with body weight, confirming the assumption that heavier animals have larger body sizes and therefore longer vocal tracts. It is unsurprising that the formants in the upper half of the frequency spectrum were on average better indicators of body size than the formants in the lower half, since the lowest formants are known to be more subject to variation in vocal tract shape, and the upper formants to reflect vocal tract length more accurately (Fant 1960). The same pattern of relations was found between formant frequencies, formant spacing and reproductive success, most probably as a simple consequence of the correlation between formant frequencies and body weight (which is itself correlated with reproductive success, see Methods).

The maximum vocal tract lengths that we estimated from fitting a simple vocal tract model to our acoustic data ranged from 60.5 cm in a 5 year old to 78.0 cm in an 11 year old and were consistent with preliminary post mortem measurements of the distance from the sternum to the incisors in stags culled on an adjacent part of the island (W. T. Fitch, D. Reby, B. Cargnelutti & K. McComb, unpublished data). They were also strongly correlated with age and, within adults, with both estimated body weight and our index of reproductive success. Our measures of formant spacing indicated that harsh roaring involves an additional extension of the vocal tract by ca. 2 cm. This extension may be caused either by the full stretching of the neck when the animal raises its head to its maximal extent, or by a further effort to bring the larynx closer to the sternum. Thus stags advertise their body size to maximum effect in these harsh calls, which are characteristic of periods of intense interactions with hinds and male-male contests, where displaying size to receivers of both sexes is likely to be of paramount importance.

Minimum frequency spacing of the formants, achieved when the vocal tract is most fully extended as the larynx

is lowered down to the sternum attachment of the sternothyroid muscle, therefore provides an honest indication of body size. Because the length of the fully extended vocal tract is physically determined by the length of the neck and the length of the head (which is itself likely to reflect body size and therefore correlates with body weight), the sternum limit acts as a morphological constraint which generates honest information in the acoustic structure of the call, independently of any cost linked to its production. This result emphasizes the potential importance of physical, anatomical constraints for the production of honest information in animal communication (Fitch 2000; Fitch & Hauser, *in press*).

Our results suggest that males should use formant frequencies as an indicator of their opponents' fighting ability during roaring contests and parallel walks (during which males often deliver harsh roars, personal observation), thereby avoiding fights with older (more experienced) and larger stags. Clutton-Brock & Albon (1979) found that adult harem holders were less likely to answer to playbacks of young stags' roars. We suggest that stags use filter characteristics (high formants) possibly combined with some source characteristics (high mean and maximum fundamental frequency) to identify subadult individuals. Furthermore, we would expect females to prefer stags whose roars have formants with lower minimum frequencies, as they are likely to be larger and have higher fitness.

Red deer stags have an unusual capacity to vary resonance frequencies over a wide range. This ability is present in other polygynous cervids (Corsican deer, wapiti and fallow deer), but is apparently absent in more primitive genera, such as roe deer, *Capreolus capreolus*, muntjac species, *Muntiacus* sp., or Chinese water deer, *Hydropotes inermis*, which are generally monogamous and territorial. In the primitive species the male's vocal repertoire is typically composed of territorial and antipredator bark-like noisy calls characterized by static formants (Oli & Jacobson 1995; Reby 1998; Reby et al. 1999b). More generally, most terrestrial mammals studied (with the exception of humans) seem to lack sophisticated vocal tract control and can modulate formants only by using their lips and lower jaw to control the shape and the opening condition of the mouth cavity (Lieberman et al. 1969; Shipley et al. 1991; Hauser et al. 1993). Red deer roaring therefore constitutes an interesting model not only for studying the evolution of honest signalling in mammals but also for examining the contribution of dynamic vocal tract resonances to the communication system of a mammal that lacks language.

Acknowledgments

We thank Tim Clutton-Brock for invaluable help and support throughout this study and for access to the long-term database on the Rum red deer stags. Steve Albon, Richard Clarke, Callan Duck, Fiona Guinness, Sean Morris, Glenn Iason, Josephine Pemberton and the Scottish Natural Heritage staff on Rum gave indispensable help with data collection and/or analyses. Chris Darwin

and Tecumseh Fitch provided insightful comments and guidance on vocal production in mammals and sound processing. D.R. was supported by Fyssen & Marie Curie fellowships. Equipment for sound recording and analysis came from grants from NERC and BBSRC (to K.M.). The fieldwork on Rum was supported by grants from NERC (to Tim Clutton-Brock).

References

- Carterette, E., Shipley, C. & Buchwald, J. 1979. Linear prediction theory of vocalization in cat and kitten. In: *Frontiers in Speech Communication Research* (Ed. by B. Lindblom & S. Ohman), pp. 245–257. New York: Academic Press.
- Chariff, R. A., Mitchell, S. & Clark, C. W. 1995. *Canary 1.12 User's Manual*. Ithaca, New York: Cornell Laboratory of Ornithology.
- Clutton-Brock, T. H. & Albon, S. D. 1979. The roaring of red deer and the evolution of honest advertising. *Behaviour*, **69**, 145–170.
- Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. 1982. *Red Deer Behavior and Ecology of Two Sexes*. Chicago: University of Chicago Press.
- Cowlishaw, G. 1996. Sexual selection and information content in gibbon song bouts. *Ethology*, **102**, 272–384.
- East, M. E. & Hofer, H. 1991. Loud calling in a female dominated society: II. Contexts and functions of whooping of spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, **42**, 651–669.
- Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fitch, W. T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, **102**, 1213–1222.
- Fitch, W. T. 2000. Skull dimensions in relation to body size in nonhuman primates: the causal bases for acoustic allometry. *Zoology*, **103**, 40–58.
- Fitch, W. T. & Hauser, M. D. *In press*. Unpacking honesty: generating and extracting information from acoustic signals. In: *Animal Communication* (Ed. by A. Megala-Simmons & A. Popper). Berlin: Springer-Verlag.
- Fitch, W. T. & Reby, D. 2001. The descended larynx is not uniquely human. *Proceedings of the Royal Society of London, Series B*, **268**, 1669–1675.
- Harrington, F. H. & Mech, L. D. 1979. Wolf howling and its role in territory maintenance. *Behaviour*, **68**, 207–249.
- Hauser, M. D. 1993. The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight and social context. *American Naturalist*, **142**, 528–542.
- Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, Massachusetts: MIT Press.
- Hauser, M. D., Evans, C. S. & Marler, P. 1993. The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations. *Animal Behaviour*, **45**, 423–433.
- Kruuk, L. E., Clutton-Brock, T. H., Slate, J., Pemberton, J. M., Brotherstone, S. & Guinness, F. E. 2000. Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences, U.S.A.*, **97**, 698–703.
- Lieberman, P., Klatt, D. H. & Wilson, W. H. 1969. Vocal tract limitations on the vowel repertoires of rhesus monkeys and other nonhuman primates. *Science*, **164**, 1185–1187.
- McComb, K. 1987. Roaring by red deer stags advances oestrus in hinds. *Nature*, **330**, 684–649.
- McComb, K. E. 1988. Roaring and reproduction in red deer *Cervus elaphus*. Ph.D. thesis, University of Cambridge.
- McComb, K. E. 1991. Female choice for high roaring rate in red deer, *Cervus elaphus*. *Animal Behaviour*, **41**, 79–88.

- McComb, K., Packer, C. & Pusey, A.** 1994. Roaring and numerical assesment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, **47**, 379–387.
- Masataka, N.** 1994. Lack of correlation between body size and frequency of vocalisations in young female Japanese macaques (*Macaca fuscata*). *Folia Primatologica*, **63**, 115–118.
- Mitani, J. C.** 1987. Territoriality and monogamy among agile gibbons (*Hylobates agilis*). *Behavioral Ecology and Sociobiology*, **20**, 265–269.
- Mitchell, B., McCowan, D. & Nicholson, I. A.** 1976. Annual cycles of body weight and condition in Scottish red deer *Cervus elaphus*. *Journal of Zoology*, **180**, 107–127.
- Morton, E. S.** 1977. On the occurrence and significance of motivation-structural rules in some birds and mammal sounds. *American Naturalist*, **111**, 855–869.
- Oli, M. K. & Jacobson, H. A.** 1995. Vocalisations of barking deer (*Muntiacus muntjac*) in Nepal. *Mammalia*, **59**, 179–186.
- Owren, M. J.** 1990. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: I. Natural calls. *Journal of Comparative Psychology*, **104**, 20–28.
- Press, W. H., Teukolsky, W. T., Vetterling, W. T. & Flannery, B. P.** 1992. *Numerical Recipes in C: the Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Reby, D.** 1998. Etude de la communication sonore chez les cervidés. Prise en compte de la dimension individuelle. Thèse de doctorat, l'Université Paul Sabatier.
- Reby, D., Joachim, J., Lauga, J., Lek, S. & Aulagnier, S.** 1998. Individuality in the groans of fallow deer (*Dama dama*) bucks. *Journal of Zoology*, **245**, 79–84.
- Reby, D., Cargnelutti, B. & Hewison, A. J. M.** 1999a. Contexts and possible functions of barking in roe deer. *Animal Behaviour*, **57**, 1121–1128.
- Reby, D., Cargnelutti, B., Joachim, J. & Aulagnier, S.** 1999b. Spectral acoustic structure of barking in roe deer: sex-, age-, and individual related variations. *Comptes Rendus de l'Académie des Sciences*, **322**, 271–279.
- Riede, T. & Fitch, W. T.** 1999. Vocal tract length and acoustics of vocalization in the domestic dog, *Canis familiaris*. *Journal of Experimental Biology*, **202**, 2859–2869.
- Shiple, C., Carterette, E. C. & Buchwald, J. S.** 1991. The effect of articulation on the acoustical structure of feline vocalization. *Journal of the Acoustical Society of America*, **89**, 902–909.
- Titze, I. R.** 1994. *Principles of Voice Production*. Englewood Cliffs, New Jersey: Prentice Hall.