

Red deer stags use formants as assessment cues during intrasexual agonistic interactions

David Reby^{1,*}, Karen McComb^{1,†}, Bruno Cargnelutti², Chris Darwin¹,
W. Tecumseh Fitch³ and Tim Clutton-Brock⁴

¹*School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK*

²*Institut de Recherche sur les Grands Mammifères, INRA, Toulouse, France*

³*School of Psychology, University of St Andrews, St Andrews KY16 9JU, UK*

⁴*Large Animal Research Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK*

While vocal tract resonances or formants are key acoustic parameters that define differences between phonemes in human speech, little is known about their function in animal communication. Here, we used playback experiments to present red deer stags with re-synthesized vocalizations in which formant frequencies were systematically altered to simulate callers of different body sizes. In response to stimuli where lower formants indicated callers with longer vocal tracts, stags were more attentive, replied with more roars and extended their vocal tracts further in these replies. Our results indicate that mammals other than humans use formants in vital vocal exchanges and can adjust their own formant frequencies in relation to those that they hear.

Keywords: red deer; vocal communication; formant frequencies; nonlinear phenomena

1. INTRODUCTION

Despite widespread interest in the evolution of language, little is known of how acoustic cues in human speech relate to those found in animal communication systems. In particular, while vocal tract resonances or ‘formants’ are the key parameters that determine differences between phonemes in human speech, it is still unclear what information they convey in animal vocalizations (Fitch 1997, 2000). It was believed that non-human mammals were limited in their abilities to modify formants because they lacked a ‘descended larynx’, a crucial adaptation that extended the human vocal tract, altering its shape and providing the necessary anatomy for producing the human vowels (Lieberman *et al.* 1969). However, recent research has demonstrated that the resting position of the larynx is low in the throat in a number of mammals other than humans (Weissengruber *et al.* 2002), and that some mammals have anatomical innovations that allow them to modify formant frequencies quite dramatically during vocalization (Fitch & Reby 2001; Reby & McComb 2003a).

Male red deer (*Cervus elaphus*) are known to possess a permanently descended larynx, which, in its resting position, is located halfway down the neck (Fitch & Reby 2001). When red deer stags roar as they gather and defend harems of females during the breeding season (Clutton-Brock & Albon 1979; McComb 1987, 1991), they retract the larynx even further down the neck, drawing it towards the sternum during each vocalization (Fitch & Reby 2001; Reby & McComb 2003a). It has been shown that the minimum formant frequencies, achieved when the larynx is against the sternum and the vocal tract fully extended, have the potential to provide

listeners with information on the body size of the caller (Reby & McComb 2003a). Signalling body size is of great importance in determining the outcome of agonistic interactions between animals (Davies & Halliday 1978; Clutton-Brock *et al.* 1979; Wagner 1992; Bradbury & Vehrencamp 1998). Here, we use playbacks of re-synthesized vocalizations to test whether formant frequencies function to provide cues to size in red deer. In particular, we examine whether stags adjust their agonistic behaviour according to the value of the formant frequencies in a simulated opponent’s roars, investing more in vocal displays when presented with formants indicative of larger opponents. Moreover, we test whether subjects adjust the acoustic structure of their own vocalizations in relation to the formant frequencies of their opponents, extending their own vocal tracts more fully to produce lower formants in response to callers whose formant frequencies are low.

2. MATERIAL AND METHODS

(a) Study population

Playback experiments were conducted on free-ranging red deer stags in the study population on the Isle of Rum, Inner Hebrides, Scotland (Clutton-Brock *et al.* 1982). In the Rum population, adult red deer stags gather and defend harems of females during the autumn breeding season, which reaches its peak during the first two weeks of October. During this period, stags roar repeatedly as they herd females and compete with other males (Clutton-Brock & Albon 1979; McComb 1987, 1991; Reby & McComb 2003a,b).

(b) Roars used for re-synthesis

We created our playback stimuli by re-synthesizing roars that had been recorded from four mature stags (CHE6, FLIN, RECC & TA24; aged between 8- and 9-years old) in

*Author for correspondence (reby@sussex.ac.uk).

† Authors contributed equally to this paper.

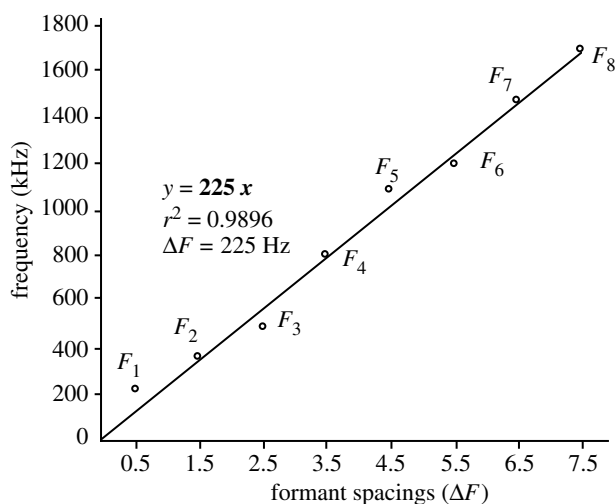


Figure 1. Illustration of the method used to estimate formants spacing and corresponding apparent vocal tract length (VTL). For each step of the analysis, the observed frequency values of each formant (F_1 – F_i) are plotted against $(2i-1)/2$ increments of the formant spacing as predicted by our vocal tract model. Then, a linear regression line is 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 ΔF for our vocal tract model. The estimated apparent VTL is finally deduced from the estimate of ΔF by applying equation (2.2) (see text). In this case, $\Delta F = 225$ Hz, which corresponds to an estimated VTL of 77.8 cm (adapted from Reby & McComb 2003a).

the Rum population during the early 1980s, using a Sennheiser MKH 816T microphone and an Uher 4200 Report Monitor open-reel recorder. Our current stags were thus unfamiliar with the calls that served as exemplars in our experiments.

For each stag exemplar, we selected a standard bout of three roars of comparable duration and spacing. The acoustic characteristics of each of these three bouts were well within the range of normal variation in the Rum population (Reby & McComb 2003a). None of these bouts contained ‘harsh roars’ (Reby & McComb 2003b), and the total duration of each bout varied between 6 and 7 s.

(c) Estimation of formant frequencies and apparent vocal tract length

Before re-synthesizing roars for playback, we extracted the formant frequencies and the overall formant frequency spacing (using the PRAAT 3.9.27 DSP package devised by P. Boersma and D. Weenink, University of Amsterdam, The Netherlands), and estimated the apparent vocal tract length (VTL) of the caller during each roar. The frequency values of the first eight formants (F_1, F_2, \dots, F_8) were extracted using linear predictive coding (LPC) via the ‘LPC: To Formants (Burg)’ command in PRAAT. Our analysis parameters were: time-step, 0.2 s; maximum number of formants, 8–10; maximum formant frequencies, 2000–2300 Hz; window of analysis, 0.1 s. The output of this analysis was transferred to EXCEL, and formant values were plotted against time and frequency and superimposed onto a narrow band spectrogram of each call. At this stage, spurious values were deleted, missing values were linearly interpolated and octave jumps were corrected for. The overall formant frequency spacing, ΔF , was then derived from the frequencies of the first

8 formants by finding the best fit for the equation

$$F_i = \frac{2i-1}{2} \Delta F, \quad (2.1)$$

which relates individual formant frequencies to average overall formant spacing in a vocal tract approximated as a straight uniform tube closed at one end (the glottis) and open at the other end (the mouth; see figure 1). Full details of this procedure are available in Reby & McComb (2003a). In the final step, we deduced the estimated apparent VTL directly from formant spacing ΔF by using the equation

$$\text{VTL} = \frac{c}{2\Delta F}, \quad (2.2)$$

where c (350 ms⁻¹) is the approximate speed of sound in a mammal vocal tract (Titze 1994; Reby & McComb 2003a).

(d) Re-synthesis of roars for playback experiments

We re-synthesized roars using a PSOLA-based algorithm that changes the apparent vocal-tract lengths for the roars by a factor of k , by modifying formant frequencies while leaving other acoustic parameters of the signal (fundamental frequency and harmonics, intensity and duration) unchanged. It operates by initially multiplying F_0 by k and duration by $1/k$ using PSOLA (Moulines & Charpentier 1990), re-sampling at the original sampling frequency multiplied by k , and, finally, playing the samples at the original sampling frequency. The script was written by C. J. Darwin, and can be downloaded from PRAAT Users Group at <http://uk.groups.yahoo.com/group/praat-users/files/Darwin%20scripts/VTchange>.

We calculated the appropriate factor (k) to alter the apparent VTLs of the original roars (estimated using the methods described in §2c) so that the re-synthesized stimuli would correspond to stags with VTLs typical of a medium subadult, small adult, medium adult and large adult in the Rum population (Reby & McComb 2003a). We created four playback stimuli that had minimum formant frequencies corresponding to extended vocal tracts of 63.2, 67, 71 and 75.8 cm, representative of medium subadult, small adult, medium adult and large adult stags, respectively.

Our acoustic modification effectively re-scales the formant frequencies in the roar stimuli for each exemplar to correspond to callers of four different VTLs, while leaving fundamental frequency and duration unchanged (figure 2). Therefore, any difference in responses shown by the deer in our playbacks can only be explained by the variation artificially imposed on the formant frequencies in the experimental stimuli. The amplitude of each of the 16 re-synthesized bouts (four size variations for each of four stag exemplars) was finally normalized and the stimuli were played back to subjects at 105 dB (peak sound pressure level (SPL) measured at 1 m from the source), which recreated SPLs typical of naturally roaring stags on Rum.

(e) Playback protocol

We conducted controlled playback experiments on free-ranging red deer stags in the Rum study population that were defending harems during the autumn breeding seasons of 2000 and 2001. Our playback protocol simulated the intrusion of a new challenger stag at close range to a harem holder. Typical responses expected from subjects are orientation and attention towards the playback source, roaring directed towards playback, and approach or retreat. Each stag

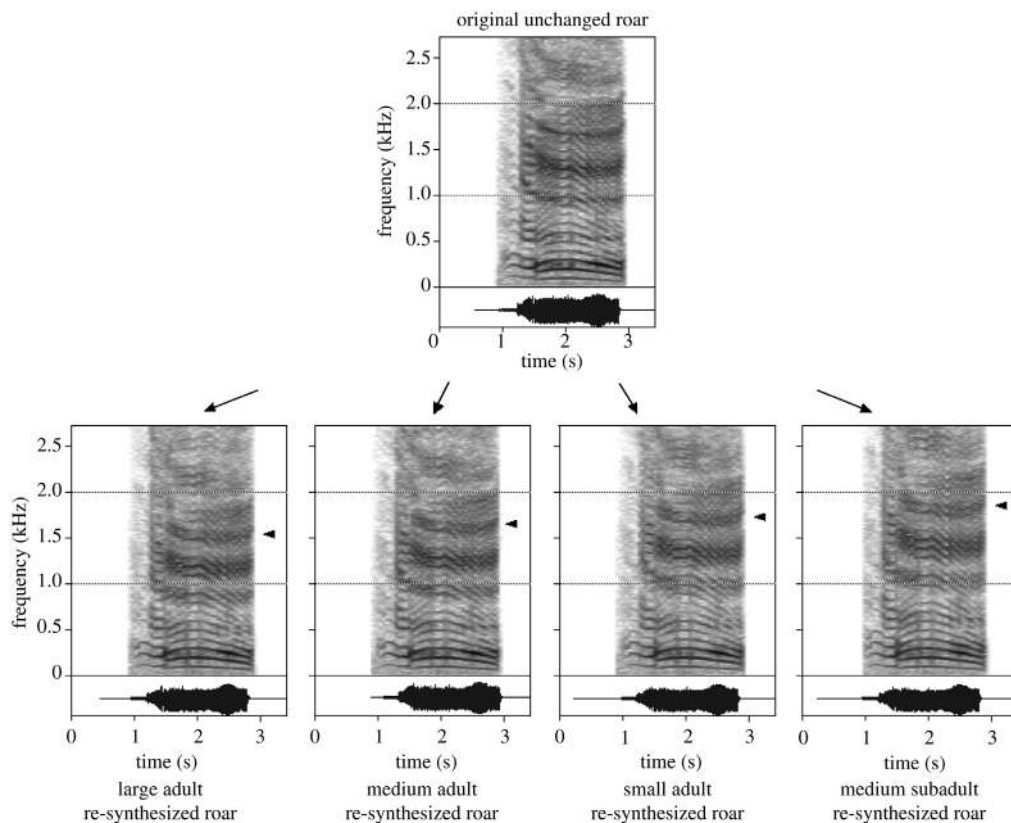


Figure 2. Re-synthesized stimuli used in playback experiments. Diagram shows spectrograms of a roar from one of the four stag exemplars, displayed as originally recorded (upper section) and in four re-synthesized roars corresponding to VTLs of 75.8, 71, 67 and 63.2 cm (lower section). The arrow indicates the position of an individual formant, which increases in frequency as apparent VTL decreases from large adult to medium subadult (full details in text).

that served as a subject in our experiments was presented with stimuli in matched pairs. A pair was a subset of stimuli from the same stag exemplar with formant frequencies representing either a medium adult paired with a medium subadult or a large adult paired with a small adult. One of the stimuli in the pair was played to the subject stag in the morning and the other in the evening of the same day, alternating whether the stimulus representing the larger or smaller size variant was presented first. Subject stags that remained actively rutting for a number of days were given a set of paired playbacks every 2 days, following the first stimulus pair by the complementary pair from the same stag exemplar and then progressing systematically through stimulus pairs from the other stag exemplars. The order of presentation of exemplars to the subject stags was randomized in a Latin square design.

Roars were played from a HHB PortaDAT through an Anchor Audio Liberty MPB4 4500H loudspeaker (set in 'music' mode; frequency response 60–15 kHz \pm 3 dB) that was positioned at approximately 100 m from the subject stag's harem (distances were verified using Leica laser rangefinder binoculars). One researcher positioned and operated the playback equipment, concealing themselves and the loudspeaker in available vegetation, while others filmed and sound-recorded the subject stag's response to playback using a Sony digital video camera (with 'fine' definition setting for image quality and 16 bit amplitude resolution, 32 kHz sampling rate for sound recording) and a Sony ECM-S959C microphone. Additional sound recordings were also made in the majority of playbacks using a Sennheiser MKH 816T microphone linked to an HHB PortaDAT recorder. Researchers carrying out the filming and sound recording positioned themselves at 50–100 m from

the subject stag's harem. Researchers used two-way radios to coordinate the timing of playback and filming. Each playback trial consisted of two presentations of a three-roar stimulus, separated by 30 s. If the stag had left his harem by the evening, or poor weather prevented playback, we played the second stimulus of the pair on the evening of the subsequent day or the day after that. Subjects were filmed for approximately 2 min before the playback started and 5 min after the onset of the playback.

(f) Analysis of behavioural responses to playback

In order to assess whether subject stags discriminated between stimuli that differed only in formant frequencies, we compared their video-recorded responses to the large versus the small variant in stimulus pairs. An edited sequence of 4 min (2 min before and 2 min after the onset of playback) was subsequently transferred to a laptop G3 PowerBook computer and stored as uncompressed DV files. Gamebreaker (SportsTec, Sydney) performance analysis software was then used to quantify stag responses in the 2 min after the onset of playback. Animal responses to the playbacks were scored by D.R. and K.M. with one frame (1/25 s) precision. From a range of behaviours measured, we summarized stag responses on the basis of three key, unambiguous behaviours: (i) the total time spent looking or walking towards the speaker (representing attention to the stimulus), (ii) the number of 'common' roars given towards playback (common roars), and (iii) the number of 'harsh' roars given towards playback (harsh roars). Common roars and harsh roars differ in acoustic structure (Reby & McComb 2003a,b; figure 3). Common roars typically sound tonal and have a spectral structure that

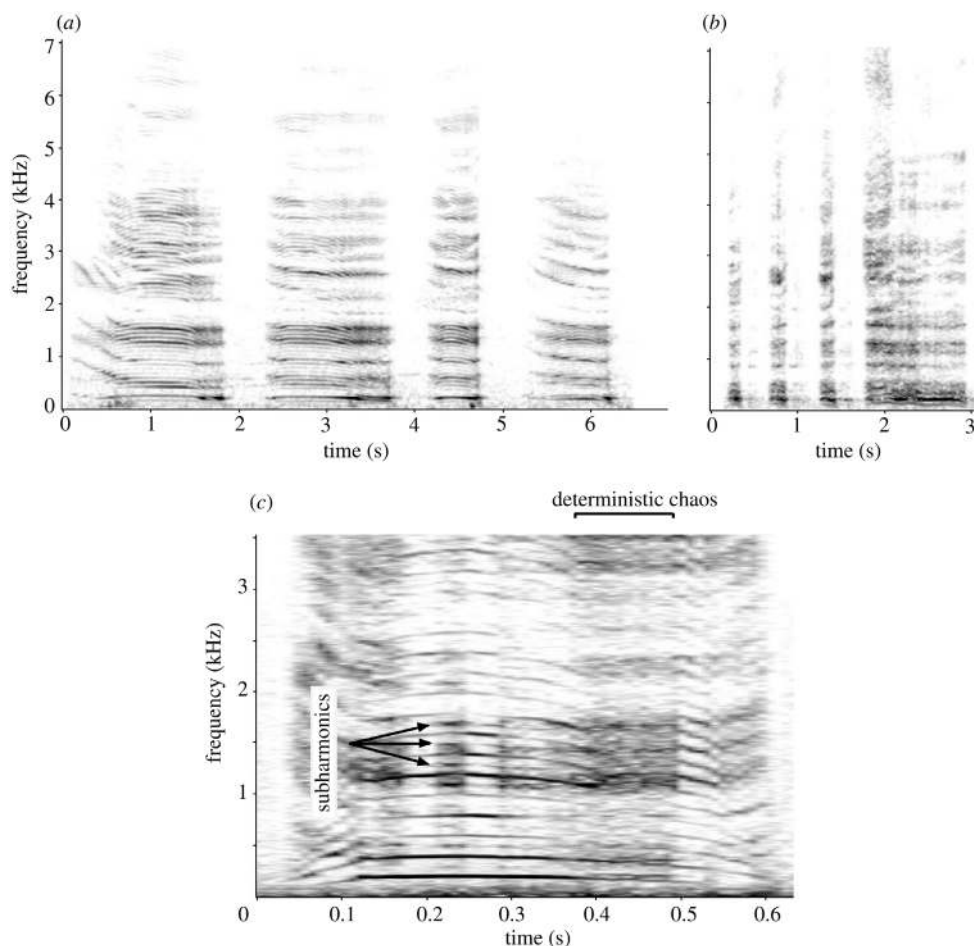


Figure 3. Narrow band spectrograms. (a) a bout of common roars, showing a clearly defined harmonic structure and modulated formants. The formants drop as the stag extends its vocal tract (by lowering its larynx) and reach a minimum plateau corresponding to the fully extended vocal tract. (b) A harsh roar, showing a noisy glottal wave (deterministic chaos) and well-defined, non-modulated formant frequencies. Formants are at their lowest throughout as harsh roars are given with a maximally extended vocal tract. (c) Detailed structure of a common roar that includes nonlinear phenomena: subharmonics and deterministic chaos; both features confer a harsh quality to roars (adapted from Reby & McComb 2003b).

shows well-defined harmonics. However, they also include noisy segments, characterized by nonlinear phenomena (Wilden *et al.* 1998; Fitch *et al.* 2002), which confer a harsh quality to the vocalization. Formants in common roars drop dramatically during the first part of the vocalization (as the stag stretches its vocal tract by lowering its larynx) and reach a minimum plateau as the vocal tract is fully extended. Harsh roars are similar to the noisiest segments of common roars (figure 3*b,c*), with the fundamental frequency and harmonics poorly defined or absent. They are also characterized by weaker formant modulation, as the stag typically fully stretches its neck and lowers its larynx before the onset of the call. Harsh roars tend to occur in situations of intense activity, after a roaring contest or during a period of repeated herding. The noisy quality of the glottal wave in harsh roars highlights the formants, which are known to reach their lowest values in these vocalizations (Reby & McComb 2003*a*).

(g) *Acoustic analysis of vocal responses to playback*

We analysed the acoustic structure of the first roar given by the subject in direct response to the playback. First, using the analysis protocol described above (§2*c*), we extracted the formant frequencies, calculated the average formant spacing and estimated the apparent VTL in the first analysable

response roar (common or harsh) given by the stag after the playback experiment. In three cases we could only accurately measure the first seven formants (F_1 – F_7) to estimate formants spacing and apparent VTL (rather than eight; see §2*c*). To control for the effect of the number of formants used on the estimation of VTL, the same number of formants was used in any paired comparison. From the formant contours, we derived the average formant spacing and the corresponding average VTL over the duration of the response roar, as well as the minimum formant spacing and the corresponding maximum apparent VTL achieved during the roar. We also noted the occurrence of nonlinear phenomena (subharmonics and deterministic chaos; see figure 3*c*) in the first response roar (common or harsh), using narrow band spectrograms in Canary 1.2.4 (filter bandwidth 44.12 Hz).

(h) *Statistical analysis*

The behavioural and acoustic measurements described in §2*f,g* were used to compare the roars given by a stag in response to a stimulus pair. Where a subject was exposed to a stimulus pair from more than one stag exemplar, the mean response for that stimulus pair was calculated across stag exemplars. Paired comparisons were performed using Wilcoxon signed ranks tests in SPSS.

Table 1. Paired comparisons (Wilcoxon tests) of behavioural and acoustical variables in playback experiments of medium adult versus medium subadult. (NLP, nonlinear phenomena. Formant spacing, from which apparent Vocal tract length (VTL) is directly calculated, is presented in square brackets.)

variable	medium adult (means \pm s.e.)		medium subadult (means \pm s.e.)		Z	n	significance
<i>behavioural variables</i>							
duration of attention (s)	70.1	5.8	48.8	8.6	-2.272	13	$p=0.0139$
number of common roars	6.31	1.58	3.10	0.72	-2.476	13	$p=0.0133$
number of harsh roars	0.69	0.17	0.23	0.12	-2.460	13	$p=0.0231$
<i>acoustic variables in first response roar</i>							
average apparent VTL (cm)	70.1	0.8	65.8	2.0	-2.668	9	$p=0.0076$
[average formant spacing (Hz)]	250	3	269	11	-2.668	9	$p=0.0076$
maximum apparent VTL (cm)	74.5	0.8	70.2	1.6	-2.668	9	$p=0.0077$
[minimum formant spacing (Hz)]	235	3	251	7	-2.668	9	$p=0.0077$
proportion of responses with NLP	0.89	0.11	0.39	0.16	-2.121	9	$p=0.0339$

Table 2. Paired comparisons (Wilcoxon tests) of behavioural and acoustical variables in playback experiments of large adult versus small adult. (NLP, nonlinear phenomena. Formant spacing, from which apparent Vocal tract length (VTL) is directly calculated, is presented in square brackets.)

variable	large adult (means \pm s.e.)		small adult (means \pm s.e.)		Z	n	significance
<i>behavioural variables</i>							
duration of attention (s)	59.3	8.4	37.7	5.2	-2.063	13	$p=0.0391$
number of common roars	4.46	1.20	2.79	0.40	-1.179	13	$p=0.2384$
number of harsh roars	0.69	0.21	0.15	0.09	-2.238	13	$p=0.0252$
<i>acoustic variables in first response roar</i>							
average apparent VTL (cm)	69.0	1.0	66.1	1.7	-1.988	10	$p=0.0469$
[average formant spacing (Hz)]	254	4	266	7	-2.191	10	$p=0.0284$
maximum apparent VTL (cm)	72.9	1.2	70.5	1.5	-2.803	10	$p=0.0051$
[minimum formant spacing (Hz)]	241	4	249	5	-2.803	10	$p=0.0051$
proportion of responses with NLP	0.72	0.13	0.22	0.13	-2.121	10	$p=0.0339$

3. RESULTS

(a) Behavioural responses to playbacks

We analysed behavioural responses in 19 pairs of playbacks simulating the medium adult versus medium subadult intruder presented to 13 different stags, and in 20 pairs of playbacks simulating the large adult versus small adult intruder presented to 13 different stags. Subjects showed significantly greater attention to the stimuli with formant frequencies representing a medium adult than to those with formants representing a medium subadult (table 1), and to those representing a large adult than to those representing a small adult (table 2). Subjects also delivered more common roars when replying to stimuli representing a medium adult than to those representing a medium subadult (table 1), but not to those representing a large adult than to those representing a small adult (table 2). Finally, subjects delivered more harsh roars when replying to stimuli representing the larger variant in both comparisons (tables 1 and 2).

(b) Acoustic parameters in vocal responses to playbacks

We were able to analyse the acoustic structure of the first response roar in 10 pairs of playbacks simulating the medium adult versus medium subadult intruder given to

9 different stags, and in 12 pairs of playbacks simulating the large adult versus small adult intruder given to 10 different stags (tables 1 and 2).

We calculated that the average formant spacing was significantly lower (by 19 Hz) in the first roar (common or harsh) given to stimuli representing the medium adult than it was in the first roar given to stimuli representing the medium subadult. This indicates that roars given to the medium adult were delivered with VTLs on average 4.3 cm longer (table 1) than those given to the medium subadult. Similarly, formants in roars given to the large adult stimuli had spacings on average 12 Hz lower than in those given to the small adult, corresponding to a 2.9 cm increase in apparent VTL (table 2).

Minimum formant spacing in both paired comparisons was also significantly lower in the response to the stimuli corresponding to the larger variant (tables 1 and 2), indicating that stags extend their vocal tract more fully when replying to stimuli with formant frequencies indicative of larger opponents.

Finally, the first roar given in response to the larger variant was more likely to contain nonlinear phenomena (subharmonics and/or deterministic chaos) in both the large adult versus small adult comparison, and the medium adult versus medium subadult comparison (tables 1 and 2).

4. DISCUSSION

Here we have demonstrated that red deer stags respond differentially to re-synthesized roars whose formant frequencies indicate stags of different sizes. Stags not only paid more attention to the roars indicating larger opponents, but they also gave more roars in response. This result is consistent with previous experiments that showed that mature harem holders roar less in response to playbacks of young stag roars than to playbacks of mature stags roars, even when roaring rate is kept constant (Clutton & Albon 1979). Furthermore, in our experiments these replies included more harsh roars: loud, conspicuous calls characterized by their noisy acoustic quality that are indicative of a high motivation level (Reby & McComb 2003b). We also found that red deer stags were able to vary their own formant frequencies by a small but consistent amount in relation to those of a perceived rival, as they tended to roar with a more fully extended vocal tract when faced with more threatening opponents. In addition, all roars that occurred as first replies tended to include more nonlinear dynamics (subharmonics and deterministic chaos), making them particularly well adapted for advertising the caller's own vocal tract resonances (Fitch *et al.* 2002).

Fitch & Reby (2001) suggested that the descended position of the larynx in red and fallow deer males may result from selection favouring individuals capable of exaggerating the impression of size conveyed by their vocalizations. Individuals with lower than average larynges yielding lower formants, would have an advantage in male–male competition and mate choice (Fitch & Reby 2001). Female red deer may also prefer stags whose roars have formants with lower minimum frequencies, as these individuals are probably larger and have higher fitness. We are currently testing these hypotheses by playing back re-synthesized roars with modified formants (mimicking stags with various VTLs) to female receivers (Charlton *et al.* in preparation). Recent research suggests that human females may also pay attention to formants in some mate choice contexts (Feinberg *et al.* 2005).

Formants are known to be of central importance in human articulated speech, their modulation forming the basis of the phonetic variation of vowels and consonants. However, their function in animal communication is still unclear. Previous studies have focused on the potential for formants to encode cues to body size (Fitch 1994, 1997; Riede & Fitch 1999; Reby & McComb 2003a,b) and context or identity (Owren *et al.* 1997; Rendall 2003), and on the ability of animals to perceive variation in formant frequencies in species-specific vocalizations (Fitch & Kelly 2000; Charlton *et al.* in preparation; Fitch & Fritz in preparation). However, whether animals actually use formants in vocal interactions had not previously been investigated. Our results indicate that, at least in one mammal species, formants play a crucial role in vocal assessment between like-sexed rivals, providing individuals with the means to assess the VTL and probable body size of opponents. Not only do red deer stags have the ability to perceive formant frequencies in an opponent's roars and alter their own behaviour accordingly, they can also adjust the acoustic structure of their own vocalizations in relation to what they hear: reserving their lowest formant frequencies for the most threatening opponents.

The importance of formants in the vocal communication of red deer and other species suggests a possible evolutionary origin for the perceptual relevance of formants in human communication (Fitch 1997): the derivation of information on body size that may be used to enhance an individual's performance when assessing rivals, prospective mates or both.

We thank Ben Charlton, Fiona Guinness, Sean Morris and Alison Donald for valuable help with fieldwork or data analysis, and the Scottish Natural Heritage for facilitating work on Rum. D.R. was supported by Fyssen and Marie Curie (EU) Fellowships. Equipment for sound recording and analysis came from grants from NERC and BBSRC (to K.M.), and the long-term research project on Rum was supported by grants from NERC (to T.C.-B.).

REFERENCES

- Bradbury, J. W. & Vehrencamp, S. L. 1998 *Principles of animal communication*. Sunderland, MA: Sinauer.
- Charlton, B. D., Reby, D. & McComb, K. In preparation. Perception of vocal tract resonances in male roars by female red deer, *Cervus elaphus*.
- 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., Albon, S. D., Gibson, R. M. & Guinness, F. E. 1979 The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Anim. Behav.* **27**, 211–225.
- Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. 1982 *Red deer: Behaviour and ecology of two sexes*. University of Chicago Press.
- Davies, N. B. & Halliday, T. R. 1978 Deep croaks and fighting assessment in toads *Bufo bufo*. *Nature* **274**, 683–685.
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M. & Perrett, D. I. 2005 Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim. Behav.* **69**, 561–568.
- Fitch, W. T. 1994 Vocal tract length perception and the evolution of language. PhD thesis, Brown University, RI.
- Fitch, W. T. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.* **102**, 1213–1222.
- Fitch, W. T. 2000 The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267.
- Fitch, W. T. & Fritz, J. B. In preparation. Primate precursors of speech perception: rhesus macaques spontaneously perceive formants in conspecific vocalizations.
- Fitch, W. T. & Kelly, J. P. 2000 Perception of vocal tract resonances by whooping cranes, *Grus americana*. *Ethology* **106**, 559–574.
- Fitch, W. T. & Reby, D. 2001 The descended larynx is not uniquely human. *Proc. R. Soc. B* **268**, 1669–1675. (doi:10.1098/rspb.2001.1704)
- Fitch, W. T., Neubaueur, J. & Herzog, H. 2002 Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav.* **63**, 407–418.
- 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**, 648–649.
- McComb, K. E. 1991 Female choice for high roaring rate in red deer, *Cervus elaphus*. *Anim. Behav.* **41**, 79–88.

- Moulines, E. & Charpentier, F. 1990 Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Commun.* **9**, 453–467.
- Owren, M. J., Seyfarth, R. M. & 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.
- Reby, D. & McComb, K. 2003a Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim. Behav.* **65**, 519–530.
- Reby, D. & McComb, K. 2003b Vocal communication and reproduction in deer. *Adv. Stud. Behav.* **33**, 231–264.
- Rendall, D. 2003 Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J. Acoust. Soc. Am.* **113**, 3390–3402.
- Riede, T. & Fitch, W. T. 1999 Vocal tract length and acoustics of vocalization in the domestic dog, *Canis familiaris*. *J. Exp. Biol.* **202**, 2859–2869.
- Titze, I. R. 1994 *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Wagner, W. E. 1992 Deceptive or honest signalling of fighting ability? A test of alternative hypotheses for the function of changes in call dominant frequency by male cricket frogs. *Anim. Behav.* **44**, 449–462.
- Weissengruber, G. E., Forstenpointner, G., Peters, G., Kübber-Heiss, A. & Fitch, W. T. 2002 Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris* f. *catus*). *J. Anat.* **201**, 195–209.
- Wilden, I., Herzog, H., Peters, G. & Tembrock, G. 1998 Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics* **9**, 171–196.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.