

The descended larynx is not uniquely human

W. Tecumseh Fitch^{1*} and David Reby^{2†}

¹Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA

²Institut de Recherche sur les Grands Mammifères, Institut National de la Recherche Agronomique (INRA), BP 27, 31326 Castanet, France

Morphological modifications of vocal anatomy are widespread among vertebrates, and the investigation of the physiological mechanisms and adaptive functions of such variants is an important focus of research into the evolution of communication. The ‘descended larynx’ of adult humans has traditionally been considered unique to our species, representing an adaptation for articulate speech, and debate concerning the position of the larynx in extinct hominids assumes that a lowered larynx is diagnostic of speech and language. Here, we use bioacoustic analyses of vocalizing animals, together with anatomical analyses of functional morphology, to document descended larynges in red and fallow deer. The resting position of the larynx in males of these species is similar to that in humans, and, during roaring, red-deer stags lower the larynx even further, to the sternum. These findings indicate that laryngeal descent is not uniquely human and has evolved at least twice in independent lineages. We suggest that laryngeal descent serves to elongate the vocal tract, allowing callers to exaggerate their perceived body size by decreasing vocal-tract resonant frequencies. Vocal-tract elongation is common in birds and is probably present in additional mammals. Size exaggeration provides a non-linguistic alternative hypothesis for the descent of the larynx in human evolution.

Keywords: laryngeal descent; cervid vocalization; formant frequencies; evolution of language; vocal-tract elongation

1. INTRODUCTION

Comparative anatomists have long been aware of a wide variety of modifications of the vertebrate vocal tract, including vocal air sacs in primates (Negus 1949; Schön Ybarra 1995), hypertrophied larynges in bats (Schneider *et al.* 1967), elongated tracheae in birds (Forbes 1882; Fitch 1999) and nasal bullae in reptiles (Martin & Bellairs 1977). An important goal of research into the evolution of communication is to understand such adaptations in terms of both proximate mechanisms and ultimate functions (Hauser 1996; Bradbury & Vehrencamp 1998). One of the best known of these adaptations is the descended larynx in our own species: adult humans are unusual in having a larynx that rests low in the throat, giving us a long pharyngeal cavity. This descended larynx has traditionally been believed to distinguish humans from all other mammals, in whom the larynx rests at the back of the oral cavity (Negus 1949; Lieberman 1984). The acoustic and evolutionary significances of this trait were first realized during the 1960s: having a low larynx allows humans to create a wider range of vocal-tract shapes, and thus more varied and distinctive speech sounds, than other mammals (Lieberman *et al.* 1969; Fitch 2000). A lowered larynx has thus been considered a key anatomical prerequisite for modern human speech, and extensive debate has focused on precisely when in hominid evolution the larynx descended (Lieberman & Crelin 1971; Falk 1975; Arensburg *et al.* 1989; Lieberman *et al.* 1989). The beliefs that a descended larynx is uniquely human and that it is diagnostic of speech have played a central role in modern theorizing about the evolution of

speech and language (Lieberman 1984; Diamond 1992; Pinker 1994; Hauser 1996; Carstairs-McCarthy 1998). In this report, we document descended larynges in males of two deer species, the red deer, *Cervus elaphus*, and the fallow deer, *Dama dama*, thus demonstrating that this feature is neither unique to *Homo sapiens* nor necessarily tied to speech production. The convergent evolution of the same trait in two independent lineages suggests a common explanation, and we offer a hypothesis that can account for laryngeal descent in both humans and deer.

During the autumn mating period or ‘rut’, red- and fallow-deer stags frequently produce loud low-pitched vocalizations (termed ‘groans’ in fallow deer and ‘roars’ in red deer). Our initial acoustic analyses of red-deer roars revealed surprisingly low-frequency resonances relative to those found in female red-deer vocalizations or in the vocalizations of other large quadrupeds (e.g. bison). Basic linear acoustics of simple tubes predicts that the frequency and spacing of vocal-tract resonances (termed ‘formants’) should be negatively correlated with vocal-tract length (Fant 1960; Titze 1994; Fitch 1997). Specifically, because $f = c/\lambda$ (where c is the speed of sound, f is the resonant frequency and λ is its wavelength), the formants of a half-open tube of constant diameter are given by:

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

where $i = 1, 2, 3, \dots$, L is the length of the tube and F_i is the frequency (in Hz) of formant i (neglecting end correction). The average frequency distance between formants, or ‘formant dispersion’, is given by

$$D_f = \frac{\sum_{i=1}^{N-1} (F_{i+1} - F_i)}{N - 1}, \quad (2)$$

*Author for correspondence (tec@wjh.harvard.edu).

†Present address: School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK.

or, by substitution, simply

$$D_f = \frac{c}{2L} \quad (3)$$

where D_f is the formant dispersion (in Hz) and N is the total number of formants measured (Fitch 1997). Note that D_f does not depend on end conditions, being valid for any tube configuration (open or closed at both ends, or half open). Thus, basic acoustics predicts an inverse correlation between vocal-tract length and formant frequencies. In empirical tests of this prediction, formant dispersion is indeed strongly and inversely correlated with anatomical vocal-tract length (distance from the larynx to the lips) in humans (Fant 1975) and in all other mammalian species examined thus far (Fitch 1997; Riede & Fitch 1999). Thus, formant values measured from vocalizations can be used to predict anatomical vocal-tract length.

Preliminary analyses of red-deer calls showed that the minimum formant dispersion values from 20 roars of 13 different red-deer males ranged from 216 Hz to 285 Hz, corresponding to vocal-tract lengths of 61–81 cm. This initially seemed impossible: the vocal tract in mammals is typically confined to the oral cavity, and in red deer this averages only 29 cm (measured from 11 male skulls, range 25–33.5 cm). Our acoustic measurements thus indicated a vocal tract two to three times as long as that expected from the anatomy. However, female calls have the expected formant values ($n = 7$, $D_f = 568$ – 661 Hz), which predict reasonable vocal-tract lengths of 26–30 cm, indicating that our recording and analysis techniques were not the problem. The low formant values in male red-deer roars suggested that these animals are somehow able to elongate their vocal tracts beyond the normal anatomical limit during roaring. By closely observing males during roaring, we noticed a consistent rearward (caudal) movement of a large ventral protuberance in the stags' throats. This protuberance is termed the 'Adam's apple' by hunters and is believed to represent the underlying larynx (Whitehead 1993). (In most stags, particularly during the rut, the neck region is obscured by a heavy beard, which may explain why previous researchers failed to report this movement.) Our observation led to the hypothesis that stags retract their larynges during roaring, thus elongating the vocal tract and accounting for the extremely low formant values we obtained.

We used the bioacoustic principles described above to test this hypothesis. If the rearward-moving protuberance is indeed the larynx, there should be a strong moment-to-moment correlation between its anatomical position and the formant values measured from the corresponding call. We tested this prediction using time-synchronized audio–video analysis of calling animals. In addition, we investigated the morphological basis of laryngeal retraction using post-mortem dissections and radiographic measurements.

2. METHODS

(a) *Behaviour: video and acoustic analyses*

We digitized videotapes of roaring red-deer males (from France, $n = 3$ individuals, and New Zealand, $n = 2$ individuals; between two and six roars each) and measured formant frequencies in the roars every 40 ms using linear prediction (Markel & Gray 1976). For audio–video recording, we used a Sony

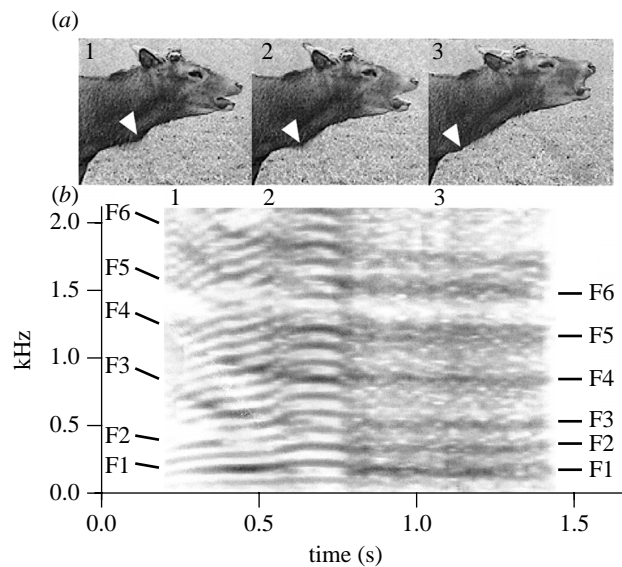


Figure 1. (a) Video of the roar of a New Zealand red deer. Landmark for frame-by-frame video measurement of vocal-tract length is indicated (the distance from the laryngeal protuberance (white triangle) to the edge of the lips). (b) Spectrogram of the emitted roar, showing decreasing formant frequencies during calling (the narrower, evenly spaced and rising frequency components in the first half of the roar represent the harmonics, which move independently of the formants due to source–filter independence).

Handycam 2006i (Sony, Tokyo, Japan) with a 100–500/5.6 Tamron zoom lens (Tamron, Commack, NY, USA) and a Telinga Pro-DAT parabolic microphone (Telinga Microphones, Tobo, Sweden). Video clips were digitized with a Media 100 system (16 bit colour at 25 frames per second, 768×576 pixels, audio tracks at 44.1 kHz, 16 bit quantization). Sound files were decimated to 5512 Hz. Formant frequencies were measured using maximum-entropy linear prediction (Press *et al.* 1992) with Praat 3.8.58 (available from Paul Boersma and David Weenink, www.praat.org). Linear prediction parameters were: 50 ms window, time step 40 ms, 8–14 formants and maximum formant frequency 3 kHz. The Praat algorithm was verified by comparison with the 'lpc' algorithm in MATLAB 5.2 with Signal Processing Toolbox (Mathworks, Inc., Natick, MA, USA). Animals were filmed at variable distances between 20 m and 90 m, so the velocity of sound created a variable audio delay; framewise audio–video synchronization was adjusted accordingly ($c = 332 \text{ m s}^{-1}$).

We measured the position of the mobile ventral protuberance (hypothesized to be the larynx) on the video images frame by frame (NIH Image 1.62, developed at the US National Institutes of Health and available on the Internet at <http://rsb.info.nih.gov/nih-image/>). Vocal-tract length was measured as the linear distance from the corner of the mouth to this protuberance. We converted this value from pixels to centimetres using each subject's total body length as a reference (figure 1). This detailed analysis was performed on a total of seven roars from the three French males.

(b) *Functional morphology: radiography and dissections*

To understand the physiological basis for vocal-tract elongation, we radiographed and dissected four adult red deer (three males) and three adult fallow deer (two males) obtained from

Table 1. Vocal-tract lengths (VTL), frequencies of the first six formants (F1–F6) and formant dispersion (D_f) in French red-deer male roars

(A: vocal-tract length at rest (pre-roar); B: minimal vocal-tract length during the roar (measured at the onset of vocalizing); C: maximal vocal-tract length during the roar; formant dispersion is calculated between the first and the sixth formants; $n = 7$ calls, three males; all values are given as mean \pm s.e.m.)

| | VTL (cm) | formant frequencies (Hz) | | | | | | D_f |
|---|----------------|--------------------------|------------------|------------------|-------------------|-------------------|-------------------|------------------|
| | | F ₁ | F ₂ | F ₃ | F ₄ | F ₅ | F ₆ | |
| A | 37.1 \pm 0.8 | — | — | — | — | — | — | — |
| B | 50.7 \pm 2.1 | 248.0 \pm 15.8 | 524.6 \pm 54.8 | 771.3 \pm 84.0 | 1283.0 \pm 80.7 | 1641.5 \pm 47.6 | 2013.0 \pm 82.3 | 353.0 \pm 14.0 |
| C | 74.3 \pm 1.7 | 194.7 \pm 9.8 | 381.1 \pm 36.4 | 526.6 \pm 22.6 | 857.8 \pm 42.5 | 1209.6 \pm 61.9 | 1510.7 \pm 42.7 | 263.2 \pm 7.2 |

deer farms or hunters in the area surrounding Toulouse, France. The location of the laryngeal protuberance was marked immediately *post mortem* and specimens were then frozen until dissection. There was no measurable distortion of the position of the larynx due to freezing or thawing. We imaged specimens radiographically at the Veterinary College of Toulouse before and after dissection. Red and fallow deer are members of the same cervid subfamily (Cervinae); as an outgroup comparison we also dissected members of a different subfamily (Odocoileinae): roe deer (*Capreolus capreolus*) and white-tailed deer (*Odocoileus virginianus*). To quantify further the sexual dimorphism in the size of the external laryngeal musculature, we weighed a 10 cm portion of the sternothyroid and sternohyoid muscles, removed 1 cm from the insertion in each specimen.

3. RESULTS

(a) Behaviour and the acoustics of roars

In red-deer stags the distance from the protuberance to the lips nearly doubled during roaring, increasing by 37% immediately before roaring and gaining another 45% during roaring (table 1). The maximum length from the protuberance to the lips averaged 75 cm (range 65–78 cm), matching closely the unusually low formant dispersion values described in § 1, which predict vocal-tract lengths of 61–81 cm.

The framewise analysis revealed a strong negative correlation between formant frequencies and the location of the protuberance ($n = 220$ frames, seven roars, 10–57 frames per roar). Formant dispersion was correlated with the length from the lips to the protuberance (vocal-tract length), with r values between -0.76 and -0.92 depending on the individual and the call analysed ($p < 0.0001$ in all cases). Vocal-tract length was also strongly correlated with individual formant values, particularly for higher formants. For example, F₆ was correlated with vocal-tract length with r values of between -0.83 and -0.93 ($p < 0.0001$), while F₄ was correlated with vocal-tract length with r values between -0.49 and -0.95 ($p < 0.015$). Overall, when all frames for all animals were combined, the correlations between the measured vocal-tract lengths and the individual formants ranged from -0.5 to -0.72 ($p < 0.0001$ in all cases). The correlations between the measured vocal-tract lengths and those estimated using the formant values and equation (3) were strong and positive ($r = 0.83$ – 0.94 , $p < 0.0001$).

These data are intelligible only if the protuberance does indeed represent the termination of the vocal tract,

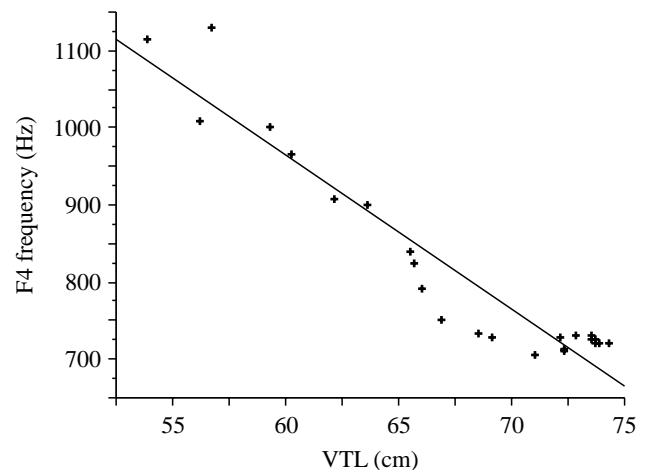


Figure 2. Effect of vocal-tract extension on formant frequencies during a red-deer roar, using the fourth formant as an example. Vocal-tract length (VTL) (measured from videos as in figure 1) is on the x -axis; fourth-formant (F₄) frequency measured from the acoustic signal is on the y -axis. Formant frequency decreases as the larynx descends.

namely the larynx (figure 2). The retraction of any other structure (e.g. muscles or glands) would have no significant effect on vocal-tract length or formant frequencies. The close correspondence between formants measured from the audio track and vocal-tract length measured from the video thus demonstrates that the larynx is pulled rearward, almost into the chest, during vocalization (figure 3). We have observed this phenomenon in red deer in southern France, northern Scotland and New Zealand, as well as in the North American red-deer subspecies (wapiti, *C. elaphus canadensis*), suggesting that retraction is not a local aberration.

(b) Radiography and dissections

Dissection confirmed that the externally visible protrusion marks the position of the underlying thyroid cartilage, thus providing a clear external indication of the larynx position in living animals, and corroborating the analysis in § 3a. Radiographs clearly indicated a low resting laryngeal position (relative to the skull base) in red- and fallow-deer males that is comparable to its low position in humans. The larynx was clearly visible in radiographs of all specimens, and its location was consistently low in the throat in male red and fallow deer

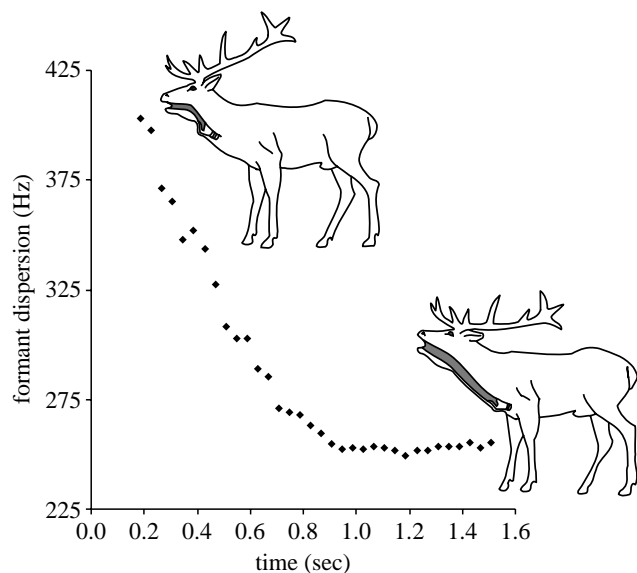


Figure 3. Retraction of the larynx during roaring in red deer leads to a sharp decrease in formant dispersion, which is negatively correlated with vocal-tract length, as the call progresses.

(figure 4), but not in females of these species nor in the roe or white-tailed deer that we examined. Dissections also revealed that male red and fallow deer possess an elongated and elastic soft palate (velum). In humans, by contrast, the velum is short and loses contact with the epiglottis as the larynx descends in early childhood (Sasaki *et al.* 1977).

The physiological basis of laryngeal retraction in males appears to be based on two components. First, the connective tissue linking the larynx to the hyoid apparatus is composed of a highly elastic thyrohyoid membrane and a loose elongated thyrohyoid muscle in red- and fallow-deer stags. In other deer, and in most mammals, this connection is made by a short tough thyrohyoid membrane, which sharply limits downward laryngeal movement relative to the hyoid and tongue. Second, the sternothyroid and sternohyoid muscles, which pull the larynx downward towards the sternum, are powerfully developed in red-deer stags, and increase in size prior to the rut (Lincoln 1971). Both muscles are highly sexually dimorphic (10 cm lengths: sternohyoid, males ($n=3$) mean 17 g (range 8–18 g) versus female 2 g; sternothyroid, males mean 20 g (range 14–30 g) versus female 8 g). Powerful laryngeal retractors and an elastic thyrohyoid linkage were both present in all adult male red and fallow deer, but not in females of these species nor in the other deer species we investigated.

4. DISCUSSION

These data show that at least two other mammalian species have a low resting larynx position comparable to humans, and that they retract the larynx even further during vocalization. Laryngeal retraction elongates the vocal tract and, thus, lowers the formant frequencies of roar and groan vocalizations. A comparison of these deer data with the laryngeal position in humans is instructive: during red-deer roaring the pharynx elongates to at least

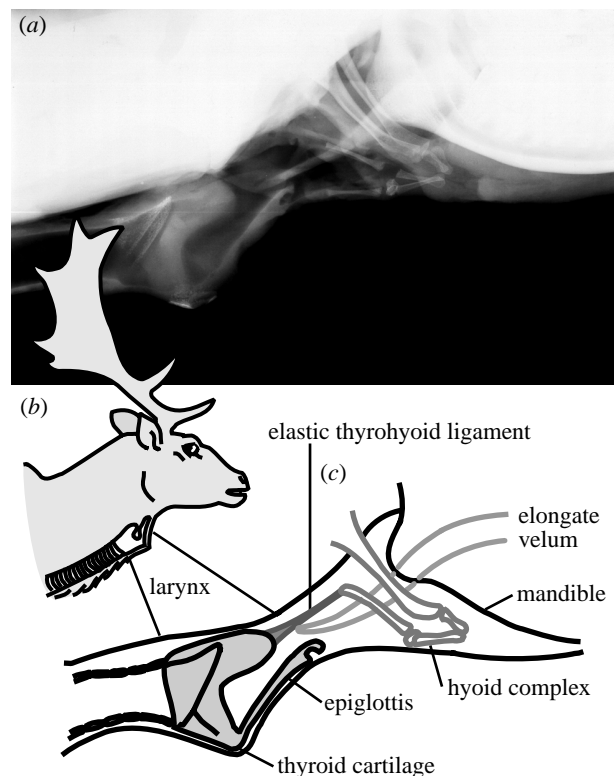


Figure 4. The resting position of the larynx in male fallow (shown) and red deer is low, as in humans. (a) Radiograph of the neck region. (b) Schematic diagram illustrating the location of the larynx. (c) Labelled outline of the radiograph. The elastic thyrohyoid ligament, which allows retraction of the larynx, is not radio-opaque and is, thus, invisible on the radiograph. Its location, as determined by dissection, is indicated.

twice the length of the oral cavity, far exceeding the 1:1 pharyngeal to oral ratio typical of adult male humans.

Although the evolutionary significance of vocal-tract elongation in deer remains unclear, it is obviously unrelated to speech and probably serves some other biological function(s). We consider two possibilities: that lowering formants aids long-distance propagation of vocalizations and that low formants serve to exaggerate the impression of size conveyed by the vocalizations. In considering these (or other) possible functions, it is important to be aware that formants are independent of voice pitch in humans, deer and all other mammals that have been investigated (Fant 1960; Lieberman & Blumstein 1988; Hauser *et al.* 1993; Fitch & Hauser 1995; Fitch 1997). Unlike wind instruments such as flutes or trumpets, where the fundamental frequency is determined by the resonances of the air column, the fundamental frequency of voices (human or otherwise) is determined by the length and tension of the vocal folds and thus by the laryngeal musculature. Formants, in contrast, reflect the length and shape of the supralaryngeal vocal tract and are independent of voice pitch (in the initial portion of figure 1b the pitch rises as the formants go down). This independence of larynx and vocal tract (source and filter) is central to modern bioacoustics and to the discussion below.

One possible function of lowered formants is to enhance low-frequency components, and thus to increase the effectiveness with which calls propagate through the

environment (Morton 1975; Michelsen 1978; Wiley & Richards 1982). In principle, because atmospheric absorption and sound scattering increase with frequency, low frequencies should propagate further than higher ones. However, when calls are emitted close to the ground (as by deer), the destructive interference between direct and ground-reflected waves increases for low frequencies, thus decreasing propagation (Piercy *et al.* 1977; Sutherland & Daigle 1997). Many studies on bioacoustic propagation have documented a peak of attenuation in the 300–800 Hz frequency range (Morton 1975; Marten & Marler 1977; Marten *et al.* 1977), which is the frequency range into which red deer lower their formants during roars. These data suggest that the formant-lowering manoeuvre documented in our study may, in fact, decrease propagation. Further difficulties for the propagation hypothesis are presented by behavioural data, which suggest that red-deer roaring serves primarily for short-range communication, especially assessment between rival males. Red deer are significantly more likely to roar when competing stags are nearby than when they are distant, and the roaring rate increases when other mature males approach (Clutton-Brock & Albon 1979). Roars may serve as signals to females in the roarer's harem, again functioning at relatively close range (Clutton-Brock & Albon 1979; McComb 1987, 1991). Similar patterns have been observed in male fallow deer, which vocalize mostly when close to females or when approached by other males (to within 30 m) (McElligot & Hayden 1999). Finally, the rutting call of the North American subspecies of red deer, called wapiti or 'elk' (*C. elaphus canadensis*), is quite different from that of red deer: wapiti produce a high-pitched (typically around 1 kHz) 'whistling' or 'bugling' (Struhsaker 1967). Our observations of bugling wapiti indicate that wapiti retract their larynges during bugling, and that formant information is clearly present and audible at short range. Over longer distances this low-frequency component becomes inaudible (Murie 1932). In this subspecies at least, the lowered formants are conveying information accessible only to nearby listeners, and must be serving some function other than increased call projection. Thus, the available data do not support the hypothesis that formant lowering serves to increase the distance over which deer vocalizations propagate.

An alternative hypothesis is that formant lowering via vocal-tract elongation serves to exaggerate the impression of size conveyed by male vocalizations. In most vertebrate species, vocal-tract length is positively correlated with body size, and thus should provide acoustic cues to body size (Fitch 1997; Riede & Fitch 1999; Fitch & Giedd 1999). Any adaptation that serves to temporarily or permanently lengthen the vocal tract could thus exaggerate such body-size cues. This hypothesis has recently been tested for tracheal elongation in birds (Fitch 1999) and found to be consistent with acoustic data and both inter- and intraspecific patterns of distribution. Because vocal-tract length is the main determinant of formant frequencies, formants potentially provide an accurate acoustic cue to body size in these species (note that a variety of birds and non-human mammals can discriminate formant frequencies (Sommers *et al.* 1992; Fitch & Kelley 2000), suggesting that these acoustic cues would be discriminable in most species). Formants thus provide

a potentially 'honest' cue to body size in many terrestrial vertebrates. However, a mechanism that enables an individual to elongate its vocal tract would allow it to duplicate the formant frequencies of, and thus to mimic, a larger animal that lacks this ability. Thus, vocal-tract elongation provides a means of exaggerating the impression of size conveyed by an individual's vocalizations.

Effective size exaggeration would be highly adaptive for red and fallow deer, since both body size and vocalizations play key roles in deer behaviour, influencing aggressive interactions and mating success (Clutton-Brock & Albon 1979; McComb 1987, 1991; Clutton-Brock *et al.* 1988; McElligot & Hayden 1999). During the rut, red and fallow-deer stags vocalize actively at night (Clutton-Brock & Albon 1979; McComb 1987, 1991; McElligot & Hayden 1999), a strategy that may enhance the efficacy of vocal exaggeration relative to visual displays. Laryngeal retraction could thus lead to more effective displays for repelling rival males (Clutton-Brock & Albon 1979), attracting females (McComb 1991) or inducing females to ovulate (McComb 1987). However, one might expect that once all males can lower the larynx, the net benefit of this adaptation would be eliminated. Interestingly, red deer pull the larynx down to the sternal attachment of the sternothyroid muscle, beyond which further retraction would be physiologically impossible. Thus, vocal-tract elongation, while creating a vocal tract nearly a third of the length of the animal's body, appears to have reached a natural physiological limit. Assuming (reasonably) that neck length is correlated with body size, the lowest formant values attained during roaring may still provide an honest indication of body size in this species, as in other mammals.

Although our data are restricted to deer, scattered anatomical data suggest that other mammals may possess a descended and/or retractable larynx. The most obvious example concerns the peculiar hyoid anatomy of the roaring cats (lions, tigers, etc., genus *Panthera*), which have long been known to possess an extremely elastic stylohyoid ligament, allowing the larynx to be retracted 20–23 cm (Owen 1834; Pocock 1916). Although this anatomical feature is well-known to taxonomists and anatomists (Hast 1989), its functional significance has long been a mystery. In koalas (*Phascolarctos cinereus*) the resting position of the larynx appears from published figures to be near the sternum (Sonntag 1921), and an elastic linkage between the skull and the larynx has been reported in other mammalian species (e.g. elephants; Gasc 1967). Although more detailed anatomical, acoustic and behavioural work is necessary, these observations suggest that other mammals may lower the larynx. Finally, at least 60 bird species possess an elongated vocal tract, which is achieved via tracheal elongation (in birds the sound source, the syrinx, lies at the base of the trachea; thus, the avian vocal tract includes the trachea, in contrast to that of mammals). A recent comparative analysis of tracheally elongated birds (Fitch 1999) supports the predictions of the size-exaggeration hypothesis for most of these species. These comparative data suggest that vocal-tract elongation is a relatively widespread response to ubiquitous constraints imposed by basic physics intersecting with the physiology of vertebrate vocal production.

The significance of these findings for human evolution and speech is two-fold. First, the fact that laryngeal descent occurs in species that lack speech clearly indicates the existence of non-phonetic functions of a lowered larynx (whatever those functions might be). This indicates that attempts to reconstruct the linguistic abilities of fossil hominids based solely on the positions of their larynges must be viewed with caution, since non-speech factors might have influenced laryngeal lowering in our lineage as well. Second, and more specifically, the size-exaggeration hypothesis that we have proposed for deer provides an explicit alternative explanation for laryngeal descent in human evolution: that laryngeal descent initially functioned to exaggerate the impression of body size conveyed by the voice, rather than to enhance the phonetic repertoire (Fitch 1997; Ohala 1983). Consistent with this second idea, contemporary human males show a sexually dimorphic lowering of the larynx at puberty. While this pubertal descent of the larynx does not increase the phonetic abilities of teenage boys, it does increase their vocal-tract lengths and lower their formant frequencies (Fant 1975; Ohala 1983; Fitch & Giedd 1999), giving adult men a more impressive baritone voice timbre. This male-specific pubertal descent of the larynx in humans thus appears to be directly analogous to the much more extreme sexually dimorphic lowering we have documented in adult male red and fallow deer. Our data suggest that a descent of the larynx serving simply to exaggerate size could have pre-dated, and perhaps served as a preadaptation for, speech- or language-specific functions of the descended larynx.

We thank Bruno Cargnelutti, Stéphane Aulagnier and George Gonzales at Institut de Recherche sur les Grands Mammifères–Institut National de la Recherche Agronomique (IRGM–INRA), Yves Lignereux, Richard Rey and Jacques Ducos de Lahitte at Ecole National Vétérinaire de Toulouse (ENVT), Geoff Asher at Invermay AgResearch and Catherine Souef of Picarel deer farm for help and access to deer. We thank Tim Clutton-Brock, Marc Hauser, Karen McComb and two anonymous reviewers for comments on earlier drafts of the manuscript. This work was supported by grants from the National Institutes of Health to W.T.F. (NIDCD T32 DC00038) and from INRA and Fondation Fyssen to D.R.

REFERENCES

- Arensburg, B., Tillier, A. M., Vandermeersch, B., Duda, H., Schepartz, L. A. & Rak, Y. 1989 A middle paleolithic human hyoid bone. *Nature* **338**, 758–760.
- Bradbury, J. W. & Vehrencamp, S. L. 1998 *Principles of animal communication*. Sunderland, MA: Sinauer Associates.
- Carstairs-McCarthy, A. 1998 Synonymy avoidance, phonology and the origin of syntax. In *Approaches to the evolution of language* (ed. J. R. Hurford, M. Studdert-Kennedy & C. Knight), pp. 279–296. New York: Cambridge University Press.
- 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., Green, D., Hiraiwa-Hasegawa, M. & Albon, S. D. 1988 Passing the buck: resource defense, lek breeding and mate choice in fallow deer. *Behav. Ecol. Sociobiol.* **23**, 281–296.
- Diamond, J. 1992 *The third chimpanzee*. New York: HarperCollins.
- Falk, D. 1975 Comparative anatomy of the larynx in man and the chimpanzee: implications for language in Neanderthal. *Am. J. Physical Anthropol.* **49**, 171–178.
- Fant, G. 1960 *Acoustic theory of speech production*. The Hague: Mouton & Co.
- Fant, G. 1975 Non-uniform vowel normalization. *Speech Trans. Lab. Q. Progr. Status Rep.* **2–3**, 1–19.
- Fitch, W. T. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoustic. Soc. Am.* **102**, 1213–1222.
- Fitch, W. T. 1999 Acoustic exaggeration of size in birds by tracheal elongation: comparative and theoretical analyses. *J. Zool. (Lond.)* **248**, 31–49.
- Fitch, W. T. 2000 The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267.
- Fitch, W. T. & Giedd, J. 1999 Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoustic. Soc. Am.* **106**, 1511–1522.
- Fitch, W. T. & Hauser, M. D. 1995 Vocal production in nonhuman primates: acoustics, physiology, and functional constraints on 'honest' advertisement. *Am. J. Primatol.* **37**, 191–219.
- Fitch, W. T. & Kelley, J. P. 2000 Perception of vocal tract resonances by whooping cranes, *Grus americana*. *Ethology* **106**, 559–574.
- Forbes, W. A. 1882 On the convoluted trachea of two species of manucode (*Manucodia atra* and *Phonygama gouldii*); with remarks on similar structures in other birds. *Proc. Zool. Soc. Lond.* **1882**, 347–353.
- Gasc, J.-P. 1967 Squelette hyobranchiale. In *Traité de Zoologie*, vol. 16 (ed. P. Grassé), pp. 550–583. Paris: Masson.
- Hast, M. 1989 The larynx of roaring and non-roaring cats. *J. Anat.* **163**, 117–121.
- Hauser, M. D. 1996 *The evolution of communication*. Cambridge, MA: MIT Press.
- Hauser, M. D., Evans, C. S. & Marler, P. 1993 The role of articulation in the production of rhesus monkey (*Macaca mulatta*) vocalizations. *Anim. Behav.* **45**, 423–433.
- Lieberman, P. 1984 *The biology and evolution of language*. Cambridge, MA: Harvard University Press.
- Lieberman, P. & Blumstein, S. E. 1988 *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press.
- Lieberman, P. & Crelin, E. S. 1971 On the speech of Neanderthal man. *Linguistic Inquiry* **2**, 203–222.
- 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.
- Lieberman, P., Laitman, J. T., Landahl, K. & Gannon, P. J. 1989 Folk physiology and talking hyoids. *Nature* **342**, 486.
- Lincoln, G. A. 1971 The seasonal reproductive changes in the red deer stag (*Cervus elaphus*). *J. Zool. (Lond.)* **163**, 105–123.
- McComb, K. 1987 Roaring by red deer stags advances date of oestrous in hinds. *Nature* **330**, 648–649.
- McComb, K. E. 1991 Female choice for high roaring rates in red deer, *Cervus elaphus*. *Anim. Behav.* **41**, 79–88.
- McElligott, A. G. & Hayden, T. J. 1999 Context related vocalisation rates of fallow bucks, *Dama dama*. *Anim. Behav.* **58**, 1095–1104.
- Markel, J. D. & Gray, A. H. 1976 *Linear prediction of speech*. New York: Springer Verlag.
- Marten, K. & Marler, P. 1977 Sound transmission and its significance for animal vocalization. 1. Temperate habitats. *Behav. Ecol. Sociobiol.* **2**, 271–290.
- Marten, K., Quine, D. B. & Marler, P. 1977 Sound transmission and its significance for animal vocalization. 2. Tropical habitats. *Behav. Ecol. Sociobiol.* **2**, 291–302.
- Martin, B. G. H. & Bellairs, A. d'A. 1977 The narial excrescence and pterygoid bulla of the gharial, *Gavialis gangeticus* (Crocodylia). *J. Zool. (Lond.)* **182**, 541–558.

- Michelsen, A. 1978 Sound reception in different environments. In *Sensory ecology: review and perspectives* (ed. M. Ali), pp. 345–373. New York: Plenum Press.
- Morton, E. S. 1975 Ecological sources of selection on avian sounds. *Am. Nat.* **109**, 17–34.
- Murie, O. J. 1932 Elk calls. *J. Mammalogy* **13**, 331–336.
- Negus, V. E. 1949 *The comparative anatomy and physiology of the larynx*. New York: Hafner Publishing Company.
- Ohala, J. J. 1983 Cross-language use of pitch: an ethological view. *Phonetica* **40**, 1–18.
- Owen, R. 1834 On the anatomy of the cheetah, *Felis jubata*. *Trans. Zool. Soc. (Lond.)* **1**, 129–136.
- Piercy, J. E., Embelton, T. F. & Sutherland, L. C. 1977 Review of noise propagation in the atmosphere. *J. Acoustic. Soc. Am.* **61**, 1403–1418.
- Pinker, S. 1994 *The language instinct*. New York: William Morrow and Company, Inc.
- Pocock, R. I. 1916 On the hyoidean apparatus of the lion (*F. leo*) and related species of Felidae. *Ann. Mag. Nat. Hist.* **8**, 222–229.
- Press, W. H., Teukolsky, W. T., Vetterling, W. T. & Flannery, B. P. 1992 *Numerical recipes in C: the art of scientific computing*, second ed. Cambridge University Press.
- Riede, T. & Fitch, W. T. 1999 Vocal tract length and acoustics of vocalization in the domestic dog *Canis familiaris*. *J. Exp. Biol.* **202**, 2859–2867.
- Sasaki, C. T., Levine, P. A., Laitman, J. T. & Crelin, E. S. 1977 Postnatal descent of the epiglottis in man. *Arch. Otolaryngol.* **103**, 169–171.
- Schneider, R., Kuhn, H.-J. & Kelemen, G. 1967 Der Larynx des männlichen *Hypsignathus monstrosus* Allen, 1861 (Pteropodidae, Megachiroptera, Mammalia). *Zeitschrift für wissenschaftliche Zoologie* **175**, 1–53.
- SchönYbarra, M. 1995 A comparative approach to the nonhuman primate vocal tract: implications for sound production. In *Current topics in primate vocal communication* (ed. E. Zimmerman & J. D. Newman), pp. 185–198. New York: Plenum Press.
- Sommers, M. S., Moody, D. B., Prosen, C. A. & Stebbins, W. C. 1992 Formant frequency discrimination by Japanese macaques (*Macaca fuscata*). *J. Acoustic. Soc. Am.* **91**, 3499–3510.
- Sonntag, C. F. 1921 The comparative anatomy of the koala (*Phascolarctos cinereus*) and vulpine phalanger (*Trichosurus vulpecula*). *Proc. Zool. Soc. Lond.* **39**, 547–577.
- Struhsaker, T. T. 1967 Behavior of elk (*Cervus canadensis*) during the rut. *Zeitschrift Tierpsychologie* **24**, 80–114.
- Sutherland, L. C. & Daigle, G. A. 1997 Atmospheric sound propagation. In *Encyclopedia of acoustics*, vol. 1 (ed. M. J. Crocker), pp. 341–365. New York: Wiley Interscience.
- Titze, I. R. 1994 *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Whitehead, G. K. 1993 *The Whitehead encyclopedia of deer*. Shrewsbury, UK: Swan Hill Press.
- Wiley, R. H. & Richards, D. G. 1982 Adaptations for acoustic communication in birds: sound propagation and signal detection. In *Acoustic communication in birds*, vol. 1 (ed. D. E. Kroodsma & E. H. Miller), pp. 131–181. New York: Academic Press.