The Company of Biologists

RESEARCH ARTICLE

Evidence of biphonation and source–filter interactions in the bugles of male North American wapiti (Cervus canadensis)

D. Reby[1],*, M. T. Wyman[1], R. Frey[2], D. Passilongo[3], J. Gilbert[4], Y. Locatelli[5] and B. D. Charlton[6]

ABSTRACT

With an average male body mass of 320 kg, the wapiti, *Cervus canadensis*, is the largest extant species of Old World deer (Cervinae). Despite this large body size, male wapiti produce whistle-like sexual calls called bugles characterised by an extremely high fundamental frequency. Investigations of the biometry and physiology of the male wapiti's relatively large larynx have so far failed to account for the production of such a high fundamental frequency. Our examination of spectrograms of male bugles suggested that the complex harmonic structure is best explained by a dual-source model (biphonation), with one source oscillating at a mean of 145 Hz (F0) and the other oscillating independently at an average of 1426 Hz (G0). A combination of anatomical investigations and acoustical modelling indicated that the F0 of male bugles is consistent with the vocal fold dimensions reported in this species, whereas the secondary, much higher source at G0 is more consistent with an aerodynamic whistle produced as air flows rapidly through a narrow supraglottic constriction. We also report a possible interaction between the higher frequency G0 and vocal tract resonances, as G0 transiently locks onto individual formants as the vocal tract is extended. We speculate that male wapiti have evolved such a dual-source phonation to advertise body size at close range (with a relatively low-frequency F0 providing a dense spectrum to highlight size-related information contained in formants) while simultaneously advertising their presence over greater distances using the very high-amplitude G0 whistle component.

KEY WORDS: Vocalisation, Mating call, Elk, Deer, Formants, Whistling

INTRODUCTION

Variation in body size explains a substantial proportion of the acoustic diversity of animal vocalisations (Morton, 1977; Ohala, 1984; Fitch and Hauser, 1995, 2002; Fletcher, 2004). Indeed, most vocal signals are produced and shaped by oscillators and/or resonators, and because larger species typically have larger oscillators or resonators, they tend to produce sounds with lower frequencies than smaller species (Hauser, 1993; Fletcher, 2004). In terrestrial mammals, the relationship between body size and the frequency characteristic of vocal signals can be predicted from the source–filter theory of vocal production (Fant, 1960; Taylor and Reby, 2010). The source–filter theory explicitly links production mechanisms to acoustic output by stating that mammal vocalisations are generated in a two-stage process, involving an oscillator as the sound source (typically the vocal folds in the larynx) and a resonator as the sound filter (the supra-laryngeal vocal tract) (Fant, 1960; Titze, 1994; Fitch, 1997; Taylor and Reby, 2010). Accordingly, because larger animals are expected to possess larger larynges with longer vocal folds and a longer vocal tract, they are also expected to produce vocalisations with lower fundamental frequency (F0) and vocal tract resonances (or formants).

Although this general rule of acoustic allometry is broadly verified across mammalian species (for example, human F0 and formants are much higher than elephant F0 and formants; McComb et al., 2003; Bachorowski and Owren, 1999), several exceptions are documented: for instance, some species have evolved anatomical innovations that enable them to produce abnormally low F0, such as the fleshy vocal pads of roaring cats (*Panthera* sp.) and Mongolian gazelle, *Procapra gutturosa* (Frey and Gebler, 2003; Titze et al., 2010); hypertrophied larynges in howler monkeys, *Alouatta* sp. (Dunn et al., 2015; Kelemen and Sade, 1960), and hammer-headed bats, *Hypsignathus monstrosus* (Bradbury, 1977); and even an additional, non-laryngeal set of vocal folds (termed 'velar vocal folds') in the koala, *Phascolarctos cinereus* (Charlton et al., 2013). Other species produce abnormally low formants for their size by extending their vocal tracts using descended and/or mobile larynges (red deer, *Cervus elaphus*, Reby and McComb, 2003; fallow deer, *Dama dama*, McElligott et al., 2006; Mongolian gazelle, *Procapra gutturosa*, Frey et al., 2008; goitred gazelle, *Gazella subgutturosa*, Frey et al., 2011; koala, Charlton et al., 2011; roaring cats, *Panthera* sp., Weissengruber et al., 2002), air sacs (black and white colobus monkey, *Colobus guereza*, Harris et al., 2006) and nasal proboscises (African elephant, *Loxodonta africana*, McComb et al., 2003; saiga, *Saiga t. tatarica*, Frey et al., 2007; elephant seals, *Mirounga leonina*, Sanvito et al., 2007). These anatomical adaptations are thought to evolve via selection pressures for individuals to lower frequency components, either to broadcast an exaggerated impression of their body size in reproductive contexts or to maximise signal propagation in the species' natural environment (koala, Charlton et al., 2011, 2013; red deer, Fitch and Reby, 2001; Reby and McComb, 2003; fallow deer, Vannoni and McElligott, 2008; bison, *Bison bison*, Wyman et al., 2012). In contrast, some animal species, such as sika deer, *Cervus nippon* (Minami and Kawamichi, 1992), appear to have evolved the ability to produce relatively higher pitched vocalisations than expected for their body size. Possibly the most extreme example of this is found in the wapiti or North American elk (*Cervus canadensis*).

With average male body masses of 225 kg for the smallest Tule wapiti, *Cervus canadensis nannodes*, 315 kg for the medium-sized Rocky Mountain wapiti, *C. c. nelsoni*, and 480 kg for the Manitoban wapiti, *C. c. manitobensis*, the North American wapiti

[1]Mammal Vocal Communication and Cognition Research Group, School of Psychology, University of Sussex, Brighton BN1 9QH, UK. [2]Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin 10315, Germany. [3]Department of Science for Nature and Environmental Resources, University of Sassari, Sassari 07100, Italy. [4]Laboratoire d'Acoustique de l'Université du Maine – UMR CNRS, le Mans 72085, France. [5]Réserve de la Haute Touche, Muséum National d'Histoire Naturelle, Obterre 36290, France. [6]School of Biology and Environmental Science, Science Centre West, University College Dublin (UCD), Belfield, Dublin 4, Ireland.

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

Journal of Experimental Biology

is the largest extant species of Old World deer (Cervinae) (Geist, 1999). Despite their large body size, male North American wapiti are known to produce extremely high-pitched, whistle-like rutting calls called bugles (Murie, 1932, 1951; Feighny et al., 2006). The frequency of the 'whistle' component of adult male bugles can reach above 2000 Hz (Feighny et al., 2006), which is in principle incompatible with the vocal fold length of adult males reported in this species (~34 mm; Riede and Titze, 2008). In addition, investigations of the biometry, physiology and biomechanical properties of the wapiti larynx, vocal folds and vocal tract have so far failed to identify any specific anatomical specialisations capable of producing such high-pitched vocalisations (Riede and Titze, 2008; Titze and Riede, 2010; Frey and Riede, 2013).

In the present study, we examined narrowband spectrograms from high-quality recordings to reveal previously undocumented complexity in the spectral composition of male wapiti bugles. We then combined several complementary approaches, including detailed acoustic analyses and modelling, as well as video and anatomical investigations, to investigate the possible acoustic sources and mechanisms involved in the production of bugles, examine their acoustic properties, and shed light on the possible functions of the different components of these calls.

MATERIALS AND METHODS
Video analyses of calling behaviour
To characterise calling behaviour (posture and movements), we examined videos of rutting males from three different populations/subspecies: 32 bugles from six males of Rocky Mountain wapiti (*Cervus canadensis nelsoni*) filmed in Yellowstone National Park, Wyoming, USA, and Jasper, Alberta, Canada, and 10 bugles from three males of Tule wapiti (*Cervus canadensis nannodes*) filmed in Point Reyes, California, USA. We documented the major observable changes in mouth and nostril movements that occurred during male bugles, and extracted still images from one video of a Rocky Mountain wapiti to illustrate how these changes coincide with the main acoustic events in a bugle. This allowed us to then predict which muscles were likely to be driving any mouth and nostril movements that were observed.

Audio recordings of male bugles
We recorded four adult male wapiti kept in captivity at two deer farms in New Zealand, including two pure-bred Manitoban wapitis aged 8 and 12 years, a 6-year-old pure-bred wapiti of unknown subspecies, and a 9-year-old wapiti (Manitoban×Rocky Mountain wapiti mix) and red deer hybrid (>90% wapiti). Breeding records or existing genetic analysis were used to ascertain the genetic status of all four animals. All vocalisations were recorded at distances of 20 to 50 m, using a Rode NTG-3 shotgun directional microphone connected to a Fostex FR-2 LE 2-Channel Compact Flash field recorder. Bugles were recorded at 32-bit amplitude resolution, 48 kHz frequency resolution, and saved as uncompressed .wav files. Bugles were extracted from the original recording sequences using Cool Edit Pro 2.0 (Syntrillium) sound editing software. A total of 42 bugles (10 or 11 recordings per individual) were included in the subsequent acoustic analyses.

Acoustic analyses
All acoustic analyses were conducted using the Praat DSP package (Boersma and Weenink, 2005). Call duration was measured directly from the waveform. Calls were then visualised and analysed using narrowband spectrograms (window length=0.03–0.1 s; time step=0.01 s; frequency step=250; frequency resolution=20 Hz;

dynamic range=70–90 dB; Gaussian window shape). Initial inspection of narrowband spectrograms revealed two clear and independently varying periodicities. The lower periodical source is hereafter referred to as F0, and the higher periodical source is referred to as G0. All non-automatically extracted variables were measured (for the entire dataset) by two independent observers (D.R. and D.P.). Agreement for these measures ranged between 94% and 100% for qualitative variables and between 96% and 100% for quantitative variables.

F0 and G0
To measure F0 variation, a pulse-detection-based pitch analysis was used (Voice Report command in Praat). A low-pass filter (upper limit set between 600 and 900 Hz) was applied to each recording before running the pulse-detection-based pitch extraction to remove most of the G0 component. This considerably improved the accuracy and efficacy of F0 detection. Extracted F0 measures were systematically checked against narrowband spectrograms and the following F0-related acoustic parameters were measured: F0mean (the average value of F0), F0min (the minimum value of F0) and F0max (the maximum value of F0). Because both pulse detection (Voice Report command) and cross-correlation (To Pitch command) automated algorithms failed to reliably detect G0 automatically (because of the presence of the lower F0), we estimated G0 parameters directly from narrowband spectrograms using the screen cursor. We measured the following G0-related parameters for each call: G0start (the onset value of G0), G0max (the maximum value of G0, generally corresponding to the value of the plateau) and G0end (the final value of G0). We also noted whether G0 tracked at least one of the formant frequencies (visually identified on spectrograms as sections where G0 stabilises within/tracks the value of a formant frequency). Finally, we also documented the presence of non-linear phenomena, including deterministic chaos, subharmonics and biphonation, from visual examinations of the narrowband spectrogram. Biphonation, the presence of two independent fundamental frequencies (or pitches; Herzel et al., 1994) produced by a single source or separate sources, is visualised on narrowband spectrograms as two independent frequencies that are not harmonically related (e.g. Wilden et al., 1998). If one source (F0) vibrates at a much lower frequency than the other (G0), because the airflow is then modulated by the frequency difference, biphonation leads to visible sidebands at linear combinations of F0 and G0 ($mG0\pm nF0$, where m and n are integers) (Nowicki and Capranica, 1986; Wilden et al., 1998; Zollinger et al., 2008). This is equivalent to considering that the lower F0 amplitude-modulates the higher frequency G0 (carrier frequency) (Gerhardt, 1998).

Minimum formant frequencies
Minimum formant frequencies were measured using the methodology described by Reby and McComb (2003). To estimate the formant frequencies at full vocal tract extension, we used narrowband spectrograms to identify the region within each bugle where the formant frequencies were minimal and stable (minimum duration of 0.2 s). We then extracted a power spectrum of this region (a spectral slice in Praat) and performed cepstral smoothing (bandwidth between 100 and 200 Hz) to derive the centre frequencies of the first eight formants. We carefully checked that the first eight formants had frequencies below that of the G0 whistle component at the point of measure, to ensure that this spectral component was not mistaken for a formant frequency. Formant values were estimated manually from spectrographic representations in a total of 10 cases when the above analysis returned formant values that

did not correspond closely to those derived from the cepstral smoothing. We then estimated formant frequency spacing, and corresponding apparent vocal tract length (aVTL), using the linear regression method of Reby and McComb (2003).

Relative amplitude of source components

The relative amplitude of source components was examined by comparing the sound pressure level (SPL; in dB) of F0 and G0, as well as the SPL of the strongest harmonic of these sources (termed domF and domG), regardless of its place in the harmonic series. We extracted an instantaneous spectrum (View Spectral Slice command in Praat) from the middle of each vocalisation and measured the SPL of F0, G0, domF and domG using the screen cursor (see Fig. S1). The relative difference in SPL between the two sources was then estimated by subtracting the SPL of F0 from that of G0 and the SPL of domF from that of domG.

Statistical analyses

All statistical analyses were conducted using R software (R Development Core Team, 2008). Inter-individual differences in the acoustic variables between the four males were tested within a multivariate ANOVA ([manova] command in 'stats' library) followed by single ANOVAs to summarise and test the contribution of each variable to the model ([summary.aov] command in 'stats' library). Correlations between duration, MeanF0 and aVTL (predictors, nested within male ID) and MaxG0 (outcome variable) were tested using linear mixed-effects models ([lme] command in 'nlme' library).

Anatomical investigations

Although the general anatomy of the wapiti larynx and the histology of its vocal fold are documented (Riede and Titze, 2008; Frey and Riede, 2013), we combined imaging and dissection techniques to provide additional anatomical details and measurements relevant to the vocal production hypotheses developed in the present study. To this effect, a 7.5-year-old captive male wapiti from Park la Haute Touche, France, that had been euthanised as part of the park's population management procedures was dissected. The specimen was frozen post mortem, stored at −18°C for 6 months, and subsequently thawed before computerised tomography (CT scan) imaging and dissection at the Institut National de la Recherche Agronomique (INRA), Tours, France (by D.R. and Y.L.). To CT scan the specimen in a 'calling posture', the vocal tract was artificially extended by maximally pulling the trachea down and fastening it to the sternum. The CT scans were performed on a SIEMENS Biograph 64 slice scanner with a 0.6 mm slice thickness on a pitch of 0.9 mm. Contiguous 0.6 mm slices were displayed with a 512×512 pixel matrix, with reconstructions performed with a 'B20s smooth' filter. All CT scans were displayed on a 512×512 pixel matrix and the field of view was varied according to the size of the wapiti's head. Because of damage affecting the base of the skull, we were only able to perform basic measurements of the vocal tract dimensions.

The excised larynx, including the pharynx, and the nasal region of the skull were re-frozen at −18°C and sent from INRA to the Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany, for dissection in water (by R.F.). This allowed us to perform detailed examinations and measurements of the most relevant anatomical structures (*sensu* Frey et al., 2007, 2008, 2011). Dissections were conducted layer-by-layer using a head loupe, and all stages of the dissection were documented using a Nikon DS70 digital photo camera (Nikon, Tokyo, Japan) and stored on a Compact Flash card. The images were subsequently downloaded to a PC and graphically

processed (Adobe Photoshop 5.5 and CS4; Adobe Systems, San Jose, CA, USA) to identify the individual components of the larynx, the soft palate, the nasal vestibulum and the remaining parts of the hyoid apparatus. Finally, to investigate the possible configuration of the laryngeal structures during phonation, the effect of the constriction of the larynx by the caudal pharyngeal constrictors was simulated by moderately fastening a plastic strap transversely around the larynx. The dimensions of the resulting narrow opening between the vocal folds and its distance from the intra-pharyngeal ostium were then measured.

RESULTS

General behaviour, posture and gesture

Examination of video recordings showed that male wapitis extend their neck and lower their larynx during vocalisations; however, the actual extent of vocal tract extension was difficult to establish because the neck mane concealed the lowest point of laryngeal retraction. Vapour expelled during vocalisations and flank movements indicated that bugles are produced during exhalation, and the larynx is rapidly pulled towards the sternum at the onset of the vocalisation, before rising again to its resting position at the end of the call. The mouth is kept open during the vocalisation, whereas the upper lips are curled upwards and the nostrils are moved backwards in a posture reminiscent of flehmen behaviour (Fig. 1). Backward nostril
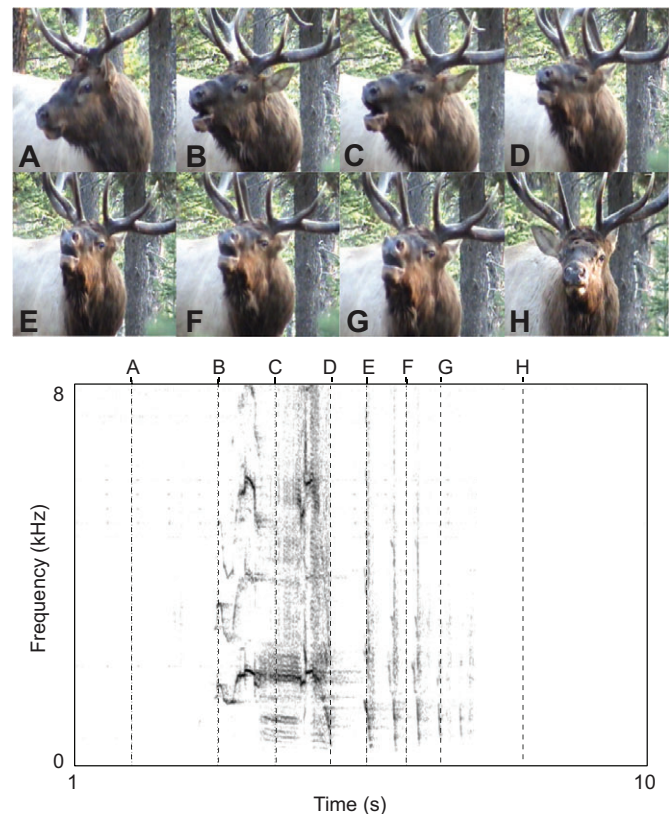

Fig. 1. Gestures during bugle production in an adult male wapiti (top) and corresponding spectrogram (bottom). The presented stages were extracted from one video clip of a rutting Rocky Mountain wapiti (*Cervus canadensis nelsoni*) in Alberta (original clip by D. Auten, 2012). The complete vocalisation consisted of one bugle of ∼2 s duration with sustained contraction of the flanks followed by three (fourth indicated) short successive 'yelps' with correspondingly short intermittent flank contractions, each lasting 0.4–0.5 s. (A) Resting position, (B) initial bugling position, (C) main bugling position, (D) final bugling position, (E) initial yelp position, (F) intermittent inhalatory position, (G) final yelp position and (H) return to resting position.

movements are closely aligned with the production of G0 during the bugle and subsequent yelps when present (Fig. 1).

Acoustics

Fundamental frequencies (F0 and G0)

Examination of narrowband spectrograms suggests that the complex harmonic structure of these vocalisations is best described as a dual-fundamental frequency signal, with a lower fundamental frequency (F0) in the 76–250 Hz range (mean±s.d.: 145.1±15.2 Hz) and a higher fundamental frequency (G0) in the 145–4187 Hz range (mean±s.d.: 1426.4±558.9 Hz; Fig. 2, Table 1). These two simultaneous and independent periodicities are clearly visible in all 42 calls. We also report sidebands at mG0±nF0 consistent with the amplitude modulation of G0 by F0 (Fig. 2). It is clear from visual examination of spectrograms from all four stags that G0 and F0 are not harmonically related (Figs 2 and 3A, and see legend of Fig. 2 for calculation of G0/F0 ratios illustrating the absence of harmonic relationship), vary independently (sometimes even in opposite directions, see Fig. 3B) and can be produced in the absence of each other (e.g. Fig. 3B). In addition, the ratio of G0max over F0max is not an integer and is highly variable both between calls (ranging between 8.7 and 26.2) and between individuals (Table 1), confirming the observation that F0 and G0 are not harmonically related and therefore correspond to biphonation. Finally, we observed that the G0 component appeared before the formant frequencies reached their minimum value in 33 out of 42 bugles, indicating that full vocal tract extension was not required for the production of G0.
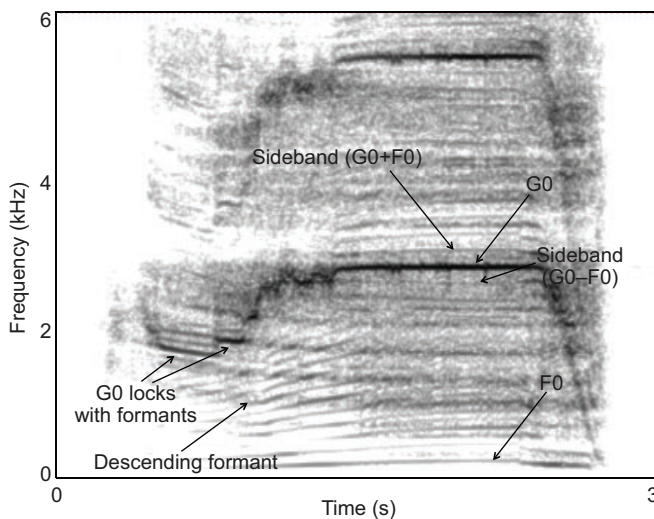

Fig. 2. Narrowband spectrogram of a typical male wapiti bugle. Two independent fundamental frequencies, a high-frequency G0 (hypothesised to be a laryngeal or a nasal aerodynamic source) and a lower-frequency F0 (hypothesised to be a laryngeal biomechanic source) are unambiguously present on the spectrogram. G0 and F0 are not harmonically related, as demonstrated by the fact that the ratio of G0/F0 is not a constant integer. This ratio changes from 1614/117 Hz=13.8 at 0.5 s into the bugle to 2707/212 Hz=12.8 at 2 s, and 1210/155 Hz=7.8 at 2.8 s. Sidebands resulting from the amplitude modulation of G0 by F0 are present at linear combinations of G0 and F0 (here visible at G0±nF0). For example, at 2 s into the call, sidebands are present at 2707−212 Hz=2495 Hz and 2707+212 Hz=2919 Hz. Neither are integer multiples of F0 and therefore they are not harmonically related to F0. The glottal source and its harmonics highlight the transfer function of the supralaryngeal vocal tract. Formants are lowered as the vocal tract is extended. Rising G0 transiently locks into successive descending formants until it reaches a maximum, modal value around 3 kHz.

Our comparison of the amplitude of the sources revealed that G0 was typically 11.3 dB higher than F0, corresponding to 2.2 times the loudness (to the human ear) and 13.5 times the acoustic intensity, as measured by median values of relative amplitude differences at the middle point of the bugles. Furthermore, the strongest harmonic of G0 (domG) was typically 7.7 dB higher than the strongest harmonic of F0 (domF), corresponding to 1.7 times the loudness (to the human ear) and 5.9 times the acoustic intensity. The strongest harmonic of the F0 source component ranged from the first to the sixteenth harmonic (median of 3rd harmonic) while the first harmonic was the strongest harmonic in the G0 source component for all calls except two. Finally, subharmonics related to the lower source F0 and deterministic chaos were present in 57% and 71% of the analysed calls, respectively.

The overall acoustic structure of bugles differed significantly between the four individuals (MANOVA, $F_{3,38}$=64, P<0.001). Protected ANOVAs revealed that all acoustic variables related to F0 and G0 differed significantly between individuals (Table 1). Linear mixed models showed that the relationships between G0max and duration, mean F0 or aVTL were not significant within individuals (duration: $F_{1,35}$=0.34, P=0.561; F0mean: $F_{1,35}$=1.54; P=0.222; aVTL: $F_{1,35}$=0.14, P=0.704).

Formant frequencies

The spectrally dense harmonic stack associated with the low F0 results in broadband excitation of the vocal tract transfer function in all bugles, so that clearly defined formant frequencies are observed. Formants drop at the beginning of the bugle and reach a minimum plateau, characteristic of the vocal tract extension already identified in red deer stags (Reby and McComb, 2003) and fallow deer bucks (McElligott et al., 2006), and consistent with the laryngeal movements observed in the videos of bugling male wapitis. Analysis of the formants from the recorded vocalisations indicates that the minimum formant spacing averages 222±7.8 Hz (mean± s.d.), corresponding to a 79±2.8 cm (mean±s.d.) maximum aVTL at full extension. Maximum aVTL ranged between 77.7 and 82.5 cm and was not significantly different between individuals ($F_{3,38}$=2.7, P=0.07).

Source–filter interactions

Examination of spectrograms also indicated that G0 is characterised by a stepwise increase in the first part of the vocalisation (termed the 'on-glide' in Feighny et al., 2006), with segments where G0 clearly 'locks' into formant frequencies: G0 transiently drops as formants decrease, until it 'jumps' to the next higher formant (see Figs 2 and 3A–C). This phenomenon was observed in all of the calls.

Anatomy

Larynx

The features inside the laryngeal cavity, including the relaxed vocal fold, are presented in Fig. 4A. Macroscopic examination of the vocal folds did not reveal any specialisations for high-frequency production, such as a vocal membrane (bats: Griffin, 1958; Suthers and Fattu, 1973; Griffiths, 1978; Hartley and Suthers, 1988; Suthers, 1988; primates: Brown and Cannito, 1995; Schön-Ybarra, 1995, see Mergell et al., 1999 for a study modelling the role of vocal membranes in phonation). The specimen had an average (left and right combined) dorsoventral vocal fold length of 34.7 mm. The maximal rostrocaudal length of the vocal fold was 10 mm, the maximal transverse width 4 mm. The laminae of the thyroid cartilage were flexible transversely. As

Table 1. Inter-individual differences for source- and filter-related variables

Male	Male 1 (N=10)	Male 2 (N=11)	Male 3 (N=11)	Male 4 (N=10)	Total (N=42)	$F_{3,38}$	P-value
Duration	2.8±0.4	3.2±0.3	2.7±0.5	2.8±0.8	2.8±0.5	1.49	0.2322
F0mean	143±13	154±8	151±18	132±11	145±13	6.19	**0.0016**
F0min	110±18	105±4	117±19	94±8	107±12	3.51	**0.0241**
F0max	177±22	215±15	193±33	186±38	193±27	5.17	**0.0042**
G0start	906±347	561±273	1974±1151	1147±429	1147±550	9.09	**0.0001**
G0max	1912±75	2716±418	3197±484	3473±296	2824±318	38.30	**<0.0001**
G0end	179±20	245±45	253±30	542±225	305±80	19.74	**<0.0001**
G0/F0	16.9±3.4	12.7±2.3	11.0±1.2	19.5±4.9	15.1±4.6	15.38	**<0.0001**
F1	174±21	166±19	169±17	158±27	167±21	1.02	0.3956
F2	312±28	274±15	324±44	279±20	296±35	7.74	**0.0004**
F3	551±56	463±66	528±46	529±43	517±62	5.39	**0.0034**
F4	698±118	627±133	686±66	647±33	663±97	1.23	0.3131
F5	953±107	933±133	959±34	952±20	949±85	0.18	0.9081
F6	1233±108	1277±109	1238±32	1267±38	1255±80	0.76	0.5251
F7	1452±82	1501±58	1442±76	1494±60	1474±71	1.91	0.1451
F8	1686±78	1703±106	1651±61	1725±55	1692±80	1.69	0.1863
ΔF	220±13	225±4	218±6	222±6	222±8	2.55	0.0694
aVTL	80±4	78±1	80±2	79±2	79±3	2.66	0.0614

F0mean, mean value of F0; F0min, minimum value of F0; F0max, maximum value of F0; G0start, onset value of G0; G0max, maximum value of G0; G0end, final value of G0; G0/F0, ratio of G0max over F0max; ΔF, overall formant frequency spacing; aVTL, apparent vocal tract length.
Data are means±s.d.

a consequence, the application of external constriction to the thyroid laminae would be expected to cause a close approximation of the arytenoids' corniculate processes and a profound narrowing of the laryngeal vestibulum.

Vocal tract

The oral vocal tract length (measured externally after skinning of the specimen as the distance from the laryngeal prominence to the lips) was 580 mm at rest and 680 mm at maximal extension when the trachea was manually pulled in a caudal direction towards the sternum. The nasal vocal tract length (measured externally after skinning of the specimen as the distance from the laryngeal prominence to the nostrils) was 600 mm at rest and 700 mm extended. The dimensions of the excised and relaxed soft palate were: 160 mm rostrocaudal length; 57 mm transverse width; and 11 mm dorsoventral thickness, halfway between the choanae and the larynx, decreasing to 7.5 mm at the rostral edge of the intra-

pharyngeal ostium. The intra-pharyngeal ostium did not display any macroscopic specialisations for high-frequency sound production. It consisted of a simple oval opening in the soft palate, measuring 45 mm in maximal rostrocaudal length and 30 mm in maximal transverse width (Fig. 4B,C). The resting length of the thyrohyoid ligament was 15 mm and could be maximally extended to 70 mm by pulling the trachea caudally towards the sternum. After excision of the hyoid apparatus and larynx, the length of the thyrohyoid ligament at maximal extension reached 120 mm. Examination of CT multi-planar reconstructions revealed that the extended oral vocal tract length was 671±2 mm and the extended nasal vocal tract length was 718±2 mm (averaged over five measures).

Nostrils

The structure of the nostrils and nasal vestibulum corresponded to that in other ruminants (see Nickel et al., 1987). The overall resting
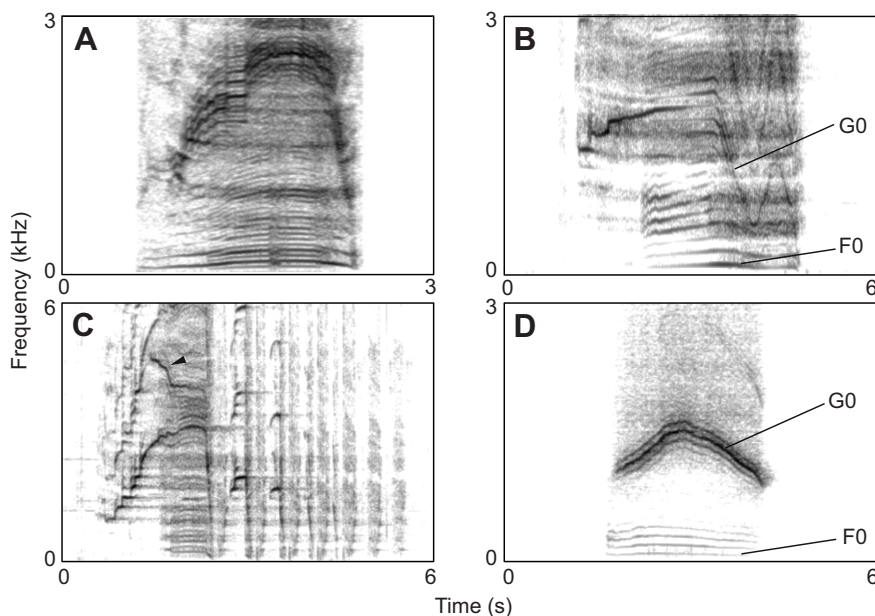

Fig. 3. Narrowband spectrograms. (A–C) Bugles from three additional adult wapitis. (A) Typical bugle with biphonation and vocal tract extension. (B) F0 starts approximately 1 s after G0, and G0 is extensively modulated independently of F0 in the last third of the bugle. (C) Arrowhead points at a possible tri-phonation episode and the bugle is followed by a series of yelps interspersed by inhalation rasps. (D) Simultaneous laryngeal voicing and lip whistling performed by D.R., illustrating how the amplitude modulation of the whistle (G0) by the voice (F0) results in sidebands similar to that observed in wapiti bugles.
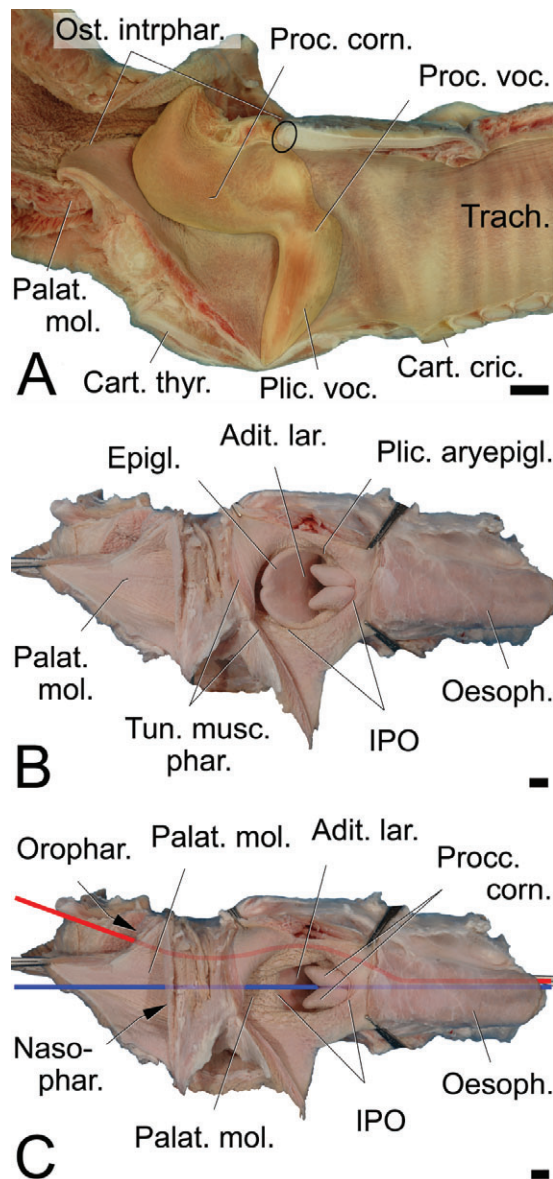
Fig. 4. Larynx and vocal fold, soft palate, intra-pharyngeal ostium and position of the laryngeal entrance. (A) The vocal fold of wapitis lacks any specialisation for high-frequency production in terms of size and structure. (B) Position during respiration and deglutition. The laryngeal entrance is pushed dorsally, through the intra-pharyngeal ostium of the soft palate into the nasal portion of the pharynx ensuring that airway and food way are separated. (C) Phonatory position. The laryngeal entrance has been retracted from the intra-pharyngeal ostium and now faces the rostral edge of the intra-pharyngeal ostium. Depending on whether the intra-pharyngeal ostium is closed by muscle contraction, the expiratory airstream can be directed either through the oral vocal tract or through both the nasal and oral vocal tract. Red line: food way; blue line: air way. (A) Mediosagittal section, medial view (original photo from Frey and Riede, 2013); (B,C) Dorsal view. Scale bars, 10 mm. Adit. lar., entrance of the larynx; Cart. cric., cricoid cartilage; Cart. thyr., thyroid cartilage; Epigl., epiglottis; IPO, intra-pharyngeal ostium; Nasophar., nasopharynx; Oesoph., oesophagus; Orophar., oropharynx; Ost. intrphar., intra-pharyngeal ostium; Palat. mol., soft palate (velum); Plic. aryepigl., aryepiglottic fold; Plic. voc., vocal fold; Proc(c). corn., corniculate process(es) of arytenoid cartilage; Proc. voc., vocal process of arytenoid cartilage; Trach., trachea; Tun. musc. phar., muscular wall of the pharynx.

anatomy of the small rostrolateral muscles acting on the nostril is summarized in Fig. 5A. The rostral ends of the muscular bundles transformed into small tendons, which further split up into smaller
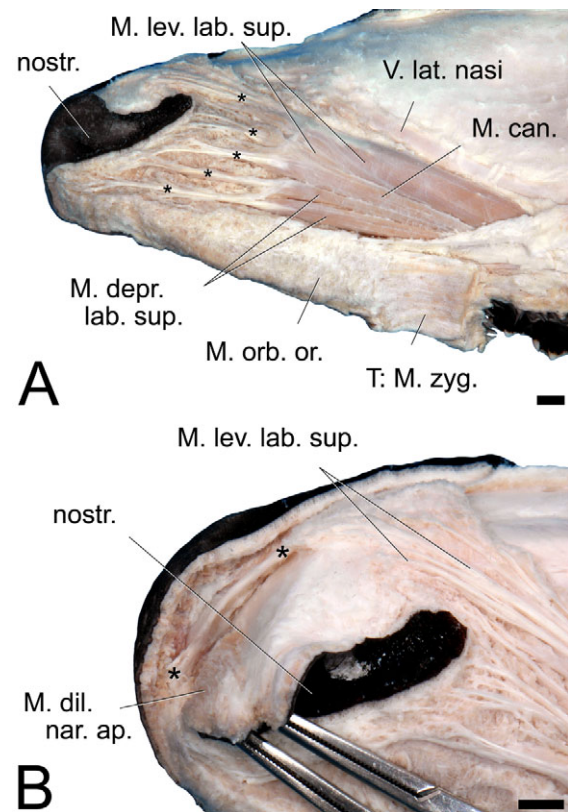

Fig. 5. Small muscles acting on the nostril and upper lip of an adult male wapiti. These muscles are decisively involved in producing the flehmen gesture and in the similar movements of the upper lip and nostrils during production of a bugle. (A) Resting position, illustrating transition of the small muscles into small tendons (asterisks) that further subdivide into delicate tendinous twigs terminating in the upper lip and around the nostril. (B) The terminal tendinous fibres of the left and right levator muscle of the upper lip unite mediorostrally to form an unpaired larger tendon that terminates in the upper lip (two asterisks). Left lateral view. Scale bars, 10 mm. M. can., caninus muscle; M. depr. lab. sup., depressor muscle of the upper lip; M. dil. nar. ap., apical dilator muscle of the nostril; M. lev. lab. sup., levator muscle of the upper lip; M. orb. or., orbicular muscle of the mouth; nostr., nostril; T: M. zyg., termination of zygomatic muscle; V. lat. nasi, lateral nasal vein.

and smaller tendinous twigs and, ultimately, ramified in the connective tissue located laterodorsally and lateroventrally to the nostrils. The small muscles coursed between the median and the lateral layer of the nasolabial levator muscle. In addition, their small tendons ran in connective tissue sheaths extending from the muscle body up to their rostral termination. The tendons of the most dorsally located small muscles, the left and right levator muscles of the upper lip, united rostromedially to form an unpaired strong tendon that terminated in the upper lip (Fig. 5B).

According to its position, course and termination, the levator muscle of the upper lip appeared to be a major relevant muscle for the flehmen-like upward curling of the nostril and nasal vestibulum we observed in the videos of calling males. Raising the dorsal edge of the nostril can be assumed to result from a concerted action of the following muscles: the nasolabial levator muscle, the levator muscle of the upper lip, the caninus muscle, the depressor muscle of the upper lip and the medial dilator muscle of the nare. Dorsocaudal pull on the upper lip, i.e. curling of the nasal vestibulum, would bring the rostral tissues closer to the fusiform alar fold, thereby creating a slit-like narrowing of this initial portion of the nasal passage.

DISCUSSION

Examination of narrowband spectrograms from high-quality recordings revealed that the bugle vocalisations of male wapitis are characterised by unusual spectral complexity. Acoustic analyses of the bugles from four adult males suggest that this complexity is best explained by the involvement of two distinct and independent sound sources (biphonation), as well as some level of source–filter interaction. We also show that the higher frequency source (G0) is characterised by much higher amplitude than the lower frequency source (F0) in calls recorded at short distance. Below we discuss the anatomical and acoustical evidence supporting these claims, as well as the possible biomechanical origins of the bugle's acoustic components. We also discuss the possible evolutionary origin of this highly unusual signal, by reviewing the likely selective advantages of its acoustic features.

Vocal tract extension

Visual inspection of spectrograms showed that the formant frequencies (excited by the low-frequency F0 source) were clearly lowered as a consequence of vocal tract extension during the initial part of the bugle. The maximum aVTL estimated from formant frequency spacing was longer, but compatible with the extended oral or nasal vocal tract length we measured in the dissected adult male. The ratio of maximum vocal tract length to body length (maxVTL/BL) in male wapitis (0.329) is similar to that of males in other species with mobile larynges and extendable vocal tracts, such as red deer (0.346), fallow deer (0.316) and goitred gazelle (0.365) (see Table 2 for details). In contrast, species that exhibit a higher resting position of the larynx and lack pronounced laryngeal mobility, as is typical for most mammals, have smaller VTL/BL ratios, e.g. sika deer (0.200), North American bison (0.147) and reindeer, *Rangifer tarandus* (0.186) (Table 2). Combined with

Table 2. Ratios of vocal tract length (VTL) to body length (BL)

Species	VTL/BL in resting state	VTL/BL in active state
Wapiti, *Cervus canadensis*	58/240=0.242[a,b]	68/240=0.283[a,b] 79/240=0.329[a,c]
Scottish red deer, *Cervus elaphus scoticus*	58/205=0.283[a,b]	68/205=0.332[a,b] 71/205=0.346[a,c]
Fallow deer, *Dama dama*	40/155=0.258[a,b]	51/155=0.329[a,b] 49/155=0.316[a,c]
Goitred gazelle, *Gazella subgutturosa*	32/126=0.254[a,b]	46/126=0.365[a,c] 43/126=0.34[a,d]
Japanese sika deer, *Cervus nippon nippon*	28/155=0.181[a,b]	31/155=0.200[a,b]
North American bison, *Bison bison*	n/a	56/380=0.147[a,c]
Reindeer, *Rangifer tarandus*	39/210=0.186[a,e]	n/a

Vocal tract lengths are provided for two states of laryngeal position: resting position (e.g. not vocalising) and active position (e.g. vocalising or manually extended), when available.
[a]Body lengths (cm) referenced from Wilson and Mittermeier, 2011.
[b]Oral VTL (cm) measured during anatomical dissections (wapiti, present study; red, fallow and sika deer, R.F., M.T.W. and D.R., unpublished data; goitred gazelle, Frey et al., 2011). Active VTL was measured during manual maximum extension of the larynx towards the sternum.
[c]Apparent VTL (cm) calculated from acoustic analysis of vocalisations (wapiti, present study; red deer, Reby and McComb, 2003; fallow deer, McElligott et al., 2006; goitred gazelle, Frey et al., 2011; bison, Wyman et al., 2012). aVTL estimation was based on apparent formant dispersion, see Reby and McComb (2003) for methodology.
[d]VTL (cm) measured from video recordings of vocalisations (goitred gazelle, Frey et al., 2011).
[e]VTL (cm) measured from CT scans of vocal tract (reindeer, Frey et al., 2007).

anatomical and video observations, these VTL/BL ratios indicate that wapiti males are extending the vocal tract during call production.

Biphonation

We identified biphonation (Neubauer et al., 2004) in all of the analysed bugles, irrespective of the subspecies and genetic profile of the four stag exemplars. Biphonation also appears to be present in the published spectrograms of 'aggressive' male bugles as well as both 'non-aggressive' and 'aggressive' female bugles recorded from free-ranging wapitis in Rocky Mountain National Park, Colorado, USA (see figs 2 and 3 from Feighny et al., 2006). In the present study, we observed two very clear and independent fundamental frequencies in male bugles: F0 ranged between 76 and 250 Hz, and G0 ranged between 145 and 4187 Hz. These two frequencies are not harmonically related and vary independently (sometimes even in opposite directions). In addition, F0 and G0 are sometimes observed separately in spectrograms, i.e. in the absence of the other, and each has its own stack of harmonic overtones (see Fig. 3A,B). Taken together, these findings strongly indicate that F0 and G0 involve distinct mechanisms of production. We were also able to identify very clear sidebands at $mG0 \pm nF0$, as predicted by the amplitude modulation of the higher G0 (carrier) by the lower F0 in the presence of biphonic sources (Gerhardt, 1998; Wilden et al., 1998). These sidebands have been mistaken for harmonics in previous literature (Feighny et al., 2006), despite the fact that they are not harmonically related to F0 (as harmonics of G0 would only be found at integer multiples of its initial frequency).

The high-pitch whistle is inconsistent with laryngeal vocal fold oscillation

Previous empirical and theoretical work on the wapiti has shown that the dimensions and histology of this species' vocal folds cannot explain the production of fundamental frequencies higher than 1.4 kHz (Riede and Titze, 2008; Titze and Riede, 2010). It has been postulated that frequencies above this could be produced by a reduction of effective vocal fold length in vibration (Riede and Titze, 2008), thereby exposing a smaller mid-membranous portion of the vocal folds to a high-impedance air flow facilitated by an unusually narrow laryngeal vestibulum (Frey and Riede, 2013). Nevertheless, because the simultaneous production of low- and high-frequency components by decoupled oscillations of a single anatomical entity, the vocal folds, is highly improbable, we suggest that laryngeal vocal fold oscillation is unlikely to explain the production of both of these spectral features. Instead, a more likely explanation is that the laryngeal vocal folds produce F0, whereas G0 is achieved either by an additional biomechanic source oscillating at a different frequency (Neubauer et al., 2004) or by an aerodynamic source.

Our anatomical investigations revealed a vocal fold length of approximately 35 mm, which accords well with previously reported male vocal fold lengths of 34 mm (Riede and Titze, 2008). These dimensions are also broadly comparable with vocal fold length in Scottish red deer (*Cervus elaphus scoticus*) stags (27 mm, Titze and Riede, 2010; 36 mm, R.F., M.T.W. and D.R., unpublished data). It is therefore not surprising that the fundamental frequencies (F0) of wapiti bugles documented in the present study are also comparable with that of European red deer stags (Scottish red deer: 107 Hz, Reby and McComb, 2003; Iberian red deer: 186 Hz, Frey et al., 2012; 181 Hz, Passilongo et al., 2013). In fact, when the high-frequency G0 is filtered out using a low-pass filter, the wapiti bugle sounds very similar to a red deer roar. Fig. 6 also illustrates how the biometric (vocal fold length) and acoustic (F0 and G0) features that we report
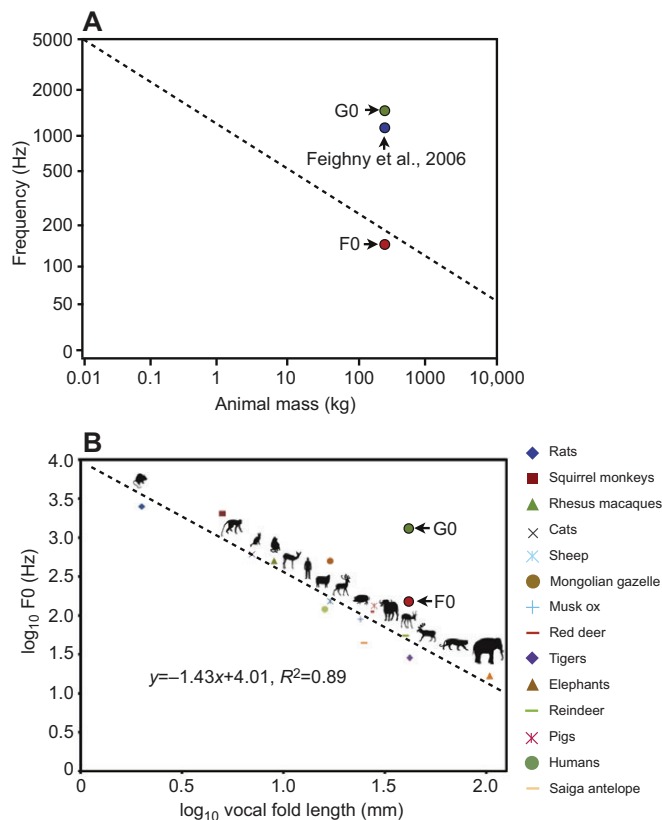
Fig. 6. Average F0 and G0 of male wapiti vocalisations. Data are shown in the context of co-variation of F0 and (A) body mass or (B) vocal fold length across terrestrial mammals.

for the bugles of male wapitis compare with that of 14 other mammal species. It is clear that while F0 follows the documented co-variation of vocal fold length and mean F0 across this sample of mammalian species, G0 is far higher than the frequency predicted for a 34 mm vocal fold. As a consequence, the extremely high G0, which reaches values over 3000 Hz in 55% of calls and even 4153 Hz in one of our recordings, is clearly inconsistent with the dimensions of the vocal folds reported for this species.

Furthermore, our dissections confirm previous studies that failed to identify any anatomical specialisation in the wapiti's upper vocal tract or larynx that could support periodic oscillations at frequencies of 2 to 4 kHz (Riede and Titze, 2008; Frey and Riede, 2013). Although the posterior edge of the velum at the level of the intra-pharyngeal ostium may act as a biomechanical valve if pressed against the dorsal wall of the nasopharynx (resembling the configuration in very high-pitched snoring; Auregan and Depollier, 1995), the relatively large dimensions of the soft palate (velum) are unlikely to support such high-frequency oscillation, unless tissue tension was extreme. Hence, we suggest that the high-frequency G0 component of wapiti bugles is more likely to be produced by an independent aerodynamic, rather than a biomechanic, source.

An aerodynamic whistle?

Two main types of aerodynamic whistles can be produced by vocal tract constrictions. (1) Vortex-induced whistles can be produced when a flow of air is forced through a narrow constriction (usually coupled with a Helmholtz resonator) such as the lips and the oral cavity in human whistling. (2) Flute-like whistles can be produced

when a flow of air forced through a narrow constriction (jet) impacts a labium (see Fabre et al., 2012). Video footage shows that vapour is exhaled at a reduced rate through the mouth during bugling before a much larger volume of vapour is expelled at the end of the vocalisation. This pattern of vapour exhalation is compatible with a reduced flow of air caused by a strong supra-laryngeal constriction during the production of the bugle that would be required to cause the high pressure necessary for the production of an aerodynamic whistle. Below, we discuss two possible production mechanisms for an aerodynamic whistle that accord with our observations in the present study.

We noticed that the occurrence and modulations of G0 were typically accompanied by characteristic nostril movements. By muscular pulling on the upper lip, i.e. dorsal curling of the nasal vestibulum, the animal could bring the rostral tissues closer to the fusiform alar fold to create a slit-like narrowing of this initial portion of the nasal passage (Fig. 7). This constriction could then support the production of a vortex-induced whistle, with the nasopharyngeal vocal tract acting as an acoustical resonator. However, given that the mammalian vocal tract comprises two nostrils, one might expect more occurrence of tri-phonation with one vocal fold-related F0 and two nostril-related G0s. While tri-phonation is clearly observed in one of the recorded calls (Fig. 3C), the inherent lateral symmetry of each nostril and associated nasal cavities, as well as the fact that effective respiration requires a high synchronisation between the left and right nostrils and the nasal vocal tracts, may explain the apparent singularity of the G0 periodicity. This apparent singularity may also be accentuated by synchronisation phenomena such as that observed between adjacent organ pipes (Abel et al., 2006).

An alternative aerodynamic source could involve a supraglottic constriction and the intra-pharyngeal ostium: when the vocal tract is extended by the muscular retraction of the larynx, the soft palate (velum) is considerably stretched longitudinally and its rostral edge could act as a labium, generating a flute-like whistle, coupled with the oral and/or nasal vocal tracts, and subsequently radiated via the oral cavity and/or nostrils (Figs 8, 9). The fundamental frequency of such whistles can be predicted as a function of the velocity of the air at the level of the constriction, the diameter of the constriction, and the distance between the constriction and the labium. Using measures from CT scans of the extended vocal tract obtained from one adult male and confirmed from dissections, we were able to estimate a whistling frequency of 3270 Hz (see Appendix for the calculation of this estimate, including the source of the various parameters entered in the model), which is of the same magnitude as the maximum (and modal) G0 frequency measured in the males in this and another published study (Feighny et al., 2006), as well as with the value of G0 (1855 Hz) measured in a bugle recorded from the specimen (M. Garcia, personal communication). Although the occurrence of triphonation in one of the recorded bugles goes against the hypothesis of a unique laryngeal whistle source, we consider that this unique occurrence is insufficient to completely exclude this hypothesis.

Although we do not provide decisive evidence that either of the above mechanisms is responsible for the observed high-frequency component of the wapiti bugle, both hypotheses constitute plausible alternatives to the unlikely involvement of laryngeal vocal fold vibration. Furthermore, examination of spectrograms shows that a clear source–filter interaction occurs in wapiti bugles (Figs 2, 3). The observed locking of G0 onto the descending formant frequencies (as the vocal tract is extended) is compatible
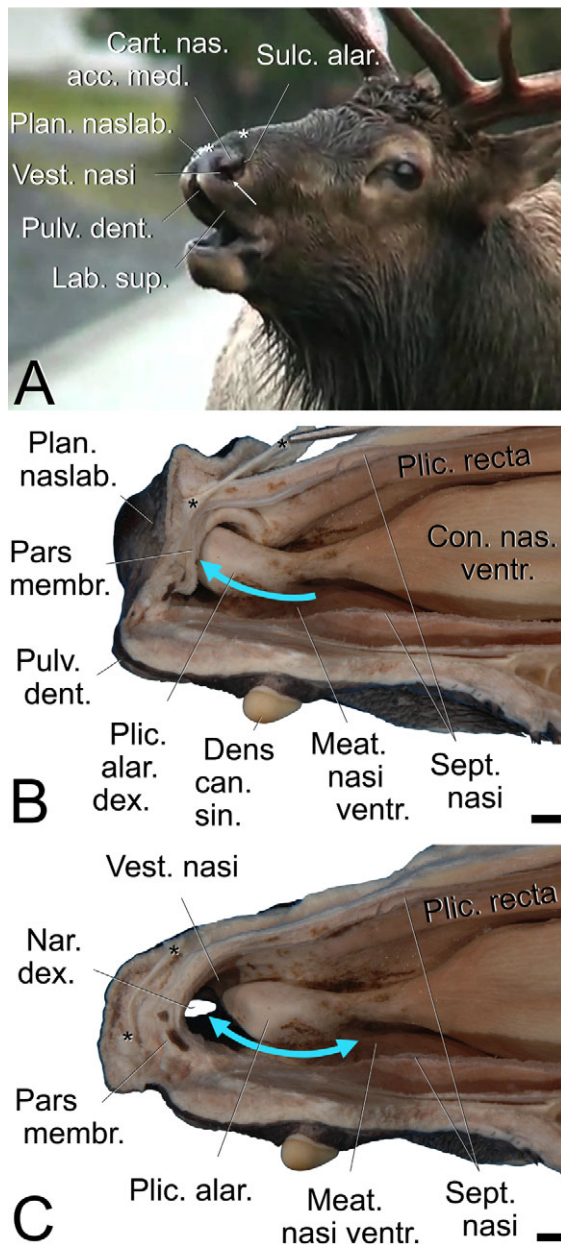
Fig. 7. Nose whistle model of G0 production. This model involves two sources, the vocal folds inside the larynx (F0 production – not shown) and the nostrils (G0 production). (A) Flehmen-like gesture of the muzzle and nostrils during production of a bugle. (B) Constriction of the rostral end of the nasal vocal tract by pulling at the tendon of the levator muscles of the upper lip. The rostral portion of the straight fold, the membranous portion of the nasal septum and the flexible nasal vestibulum form a sort of funnel into which the conical rostral tip of the alar fold tightly fits. Thereby, the connection between the nasal cavity and the nasal vestibulum is pronouncedly narrowed to an obliquely oriented slit. This constriction might be the site of a high-frequency, vortex-induced whistle (G0 production) involving the nasopharyngeal vocal tract as an acoustical resonator. (C) Upper lip, nasal vestibulum and nostril relaxed as during quiet breathing. Release of the constriction allows unrestricted inspiration and expiration along the alar fold, which is supported by the medial accessory cartilage of the nose. The asterisks indicate the position of the unpaired rostral tendon of the levator muscles of the upper lip in A–C. The small white arrow in A points to the narrow, slit-like rostral opening of the ventral nasal duct. The one-headed blue arrow in B indicates restrained phonatory expiration against the resistance of the constricted rostral end of the nasal vocal tract. The two-headed blue arrow in C indicates unrestricted inspiration and expiration through the widely open nasal vestibulum and nostril. Scale bars, 10 mm. Cart. nas. acc. med., medial accessory cartilage of the nose; Con. nas. ventr., ventral nasal concha; Dens can. sin., left upper canine tooth; Lab. sup., upper lip; Meat. nasi ventr., ventral nasal meatus; Nar. dex., right nostril; Pars membr., membranous part of nasal septum; Plan. naslab., hairless muzzle; Plic. alar. (dex.), (right) alar fold; Plic. recta, straight fold, from dorsal concha to nostril; Pulv. dent., dental pad at rostral end of hard palate; Sept. nasi, nasal septum; Sulc. alar., alar groove; Vest. nasi, nasal vestibulum.

Function and evolution

Previous functional interpretations of the extremely high G0 of male bugles have focused on increased glottal efficiency: at given laryngeal size, high G0 may optimise glottal efficiency as well as acoustic radiation through a small orifice (see Titze and Riede, 2010). Whistle elements within bugles may also function to encode quality if whistle pitch correlates with the caller's size or strength and is held reliable by a physical constraint or production cost. For example, F0 is expected to increase directly with flow velocity in both vortex and flute-like whistles (Terrien et al., 2013). As a consequence, male wapitis that are capable of producing higher G0 may advertise stronger muscles or higher lung capacities to receivers, allowing G0 to function as an index of physical quality, condition or motivational state in inter- and intra-selection contexts. Indeed, the value and duration of G0 may combine to advertise a given male's current physical condition through his ability to regularly sustain a long and high-pitched whistle. Biometric and acoustic investigations in wild ranging populations, as well as playback experiments in inter- and intra-sexual selection contexts, are now required to determine the information content of the G0 component of male bugles, and the functional relevance of any potential information.

Although high frequencies tend to propagate less efficiently than low frequencies in uniform media (Wiley and Richards, 1978), the very high-frequency G0 component is produced with a considerably higher amplitude than the glottal source F0, and may be more resilient to ground interference or low-frequency environmental noise. This might explain why F0-related components went largely unnoticed by previous researchers, and why only the whistle is audible at medium and long distances (D.R. and R.F., unpublished data). Propagation experiments in a wide range of habitats (in terms of topography and vegetation) and atmospheric conditions are now required to contrast the propagation of F0- versus G0-related components and information.

High-pitched sexual calls have evolved independently in at least one other species of Cervinae: the sika deer. However, the high-

with a feedback mechanism of the vocal tract on the source generating the whistle. Indeed, while some animals (including humans) appear to tune F0 onto a formant, or vice versa (e.g. ring doves: Riede et al., 2004; songbirds: Riede et al., 2006; gibbons: Koda et al., 2012; sopranos: Sundberg, 1975, 1979), the clear modulation of a source by a resonance, as observed here, has not been previously reported for putative biomechanic laryngeal sources (despite being predicted by simulations using a model of vocal fold tissue vibration based on morphological and biomechanical features of the wapiti vocal apparatus, if the impedance of the source and the impedance of the vocal tract were comparable; Titze and Riede, 2010). However, strong coupling is widely expected in aerodynamic sources (Fabre et al., 2012), lending further credence to the hypothesis that G0 represents an aerodynamic, rather than biomechanic, source. Interestingly, the locking/jumping of G0 onto specific resonances is also highly reminiscent of the behaviour of flute-like musical instruments (Fabre and Hirschberg, 2000).
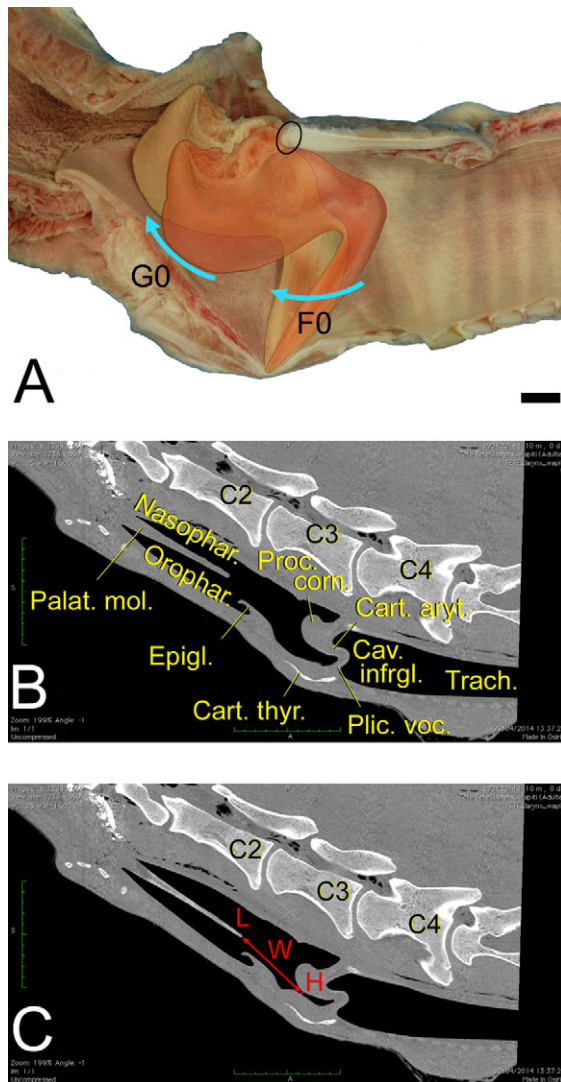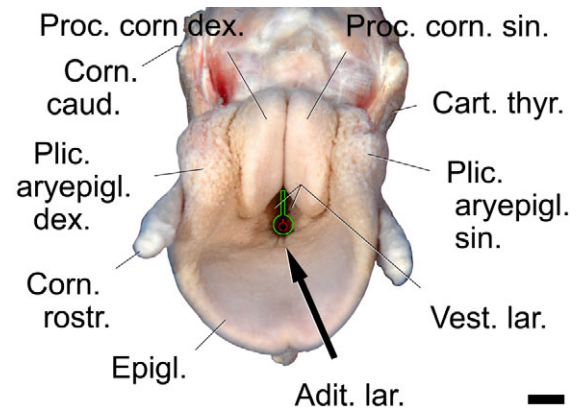
Fig. 9. Simulated constriction of the laryngeal vestibulum (either by manual compression or by fixing a plastic strap around the larynx caudal to the corniculate processes) to illustrate air-jet production in the flute model. In the live animal, this constriction would be brought about by contraction of the caudal constrictor muscles of the pharynx. Green: submaximal compression produces a ventral opening (~5 mm in diameter) plus a dorsally connected vertical slit between the walls of the vestibulum. Red: maximal compression closes the slit and leaves only a small ventral opening (~2 mm in diameter). Rostrodorsal view, compression by plastic strap (not visible). Scale bar, 10 mm. Adit. lar., entrance of the larynx; Cart. thyr., thyroid cartilage; Corn. caud. (rostr.), caudal (rostral) horn of thyroid cartilage; Epigl., epiglottis; Plic. aryepigl. dex. (sin.), right (left) aryepiglottic fold; Proc. corn. dex. (sin.), right (left) corniculate process of arytenoid cartilage; Vest. lar., laryngeal vestibulum.

196 to 1196 Hz; Minami and Kawamichi, 1992) relative to their body size. Accordingly, adult male sika deer have much smaller larynges and shorter vocal folds (~17 mm; R.F., M.T.W. and D.R., unpublished data) that are also very small relative to their body dimensions. In comparison with wapitis, sika moans and howls provide an interesting example of convergent evolution of high-frequency male sexual calls in closely related species.

Evolving a high-amplitude, high-frequency F0 is at the expense of producing spectrally dense calls that highlight vocal tract resonances (and where F0 may be an index of male androgens and associated phenotypical traits). In fact, the high-frequency portions of sexually selected sika deer male moans do not produce defined formant frequencies and sika deer do not appear to have descended larynges or extendible vocal tracts (R.F., M.T.W. and D.R., unpublished data), possibly because the very high F0 of their calls precludes the production of clear formant frequencies. In contrast, male wapitis appear to have evolved a dual mechanism to escape this constraint: we speculate that wapitis have evolved biphonation as a means of simultaneously producing a high-pitched, powerful G0 component, which may function for advertising presence to females or males over medium to long distances, while retaining a low F0 component, providing a dense spectrum to clearly highlight formants and broadcast size information at close range.

Feighny et al. (2006) report that bugles from free-ranging male wapitis only contain a low-frequency component in 'aggressive' contexts, supporting the contention that the high spectral density arising from the low F0 vocal fold source functions to facilitate the expression of body size in close-range male interactions, by highlighting vocal tract resonances. Interestingly, in another subspecies of red deer, the Iberian red deer, males give two distinct types of roars (Frey et al., 2012): short, low-pitched and low-amplitude common roars, which highlight formant frequencies and therefore efficiently communicate size-related information at short

Fig. 8. Flute model of G0 production. (A) Overlay of resting position and phonatory position of the right vocal fold to illustrate the two sources in the flute model. F0 is produced by oscillation of the vocal folds. Prerequisite for production of G0 is a narrowing of the supraglottic airway so that an air jet can be formed. This occurs when the arytenoid cartilages are rocked to tense the vocal folds for F0 production. Narrowing is enhanced by simultaneous contraction of the caudal constrictor muscles of the pharynx that press the laminae of the thyroid cartilage together, thereby compressing the laryngeal vestibulum and decreasing the diameter of the passing air jet (G0 production). Sagittal section, right half, medial view. The black oval indicates the cricoarytenoid articulation. Scale bar, 10 mm (original photo from Frey and Riede, 2013). (B,C) Virtual sagittal section of pharynx and larynx of an adult male wapiti. (B) Unmodified CT image illustrating the resting position of relevant structures for F0 and G0 production. (C) Photomontage to illustrate extension of the soft palate and supraglottal constriction of the laryngeal vestibulum. H represents the diameter of the air jet and L is the labium formed by the caudal end of the extended tensed soft palate. W is the distance between H and L. Scale bars in B and C, 100 mm. C2–4, second, third and fourth cervical vertebrae; Cart. aryt, arytenoid cartilage; Cart. thyr, thyroid cartilage; Cav. infrgl., infraglottic cavity; Epigl., epiglottis; Nasophar., nasopharynx; Orophar., oropharynx; Palat. mol., soft palate (velum); Plic. voc., vocal fold; Proc. corn., corniculate process of arytenoid cartilage; Trach., trachea.

pitched howls and moans of sika deer males are not produced by an additional source as in the biphonation of wapitis, but by the vocal folds (Herbst et al., 2013). As a consequence, sika deer calls lack a second G0 source but, instead, have a very high F0 (ranging from

distance; and longer, higher-pitched and higher-amplitude common roars, which may advertise stamina and presence over longer distances (Passilongo et al., 2013). This illustrates how different species have evolved different solutions to the challenge of producing signals that are characterised by a dense spectrum (low F0) and a high amplitude (high G0) with the relatively small larynges of terrestrial mammals. The emergence of whistle languages in several human cultures, where whistling takes over formant-based communication when communicating over long distances in mountainous environments, also provides a very interesting convergence (Classe, 1957; Rialland, 2005; Meyer, 2008). We suggest that further studies should investigate which evolutionary pressures and environmental constraints have led to the emergence of biphonation (e.g. wild dogs *Lycaon pictus*: Wilden et al., 1998; dholes: Volodina et al., 2006; horses: Briefer et al., 2015) or whistling (e.g. human whistled languages) in vertebrate vocal communication systems.

Appendix
Are the dimensions of the wapiti's vocal apparatus compatible with the production of a flute-like aerodynamic whistle?

The high-frequency whistle component of the wapiti bugle was modelled as an acoustical source produced by a flute-like instrument. Flutes, recorders and flue organ pipes are wind instruments in which the source is produced by a thin air jet (thickness H) formed by blowing through a flue channel or a slit. The jet flows across an opening in the resonator (called the mouth) toward a sharp edge (called the labium). During steady oscillations, the acoustic standing wave in the resonator drives an air flux through the mouth of the instrument, perpendicular to the jet. The acoustic velocity perturbations at the exit of the flue channel induce a modulation of the vorticity in the shear layers delimiting the jet. The vorticity perturbation is amplified by hydrodynamic instability as it is convected toward the labium. The oscillation of the jet around the labium induces an unsteady aerodynamic force. The reaction force of the labium on the air is the source of sound driving the acoustic resonator oscillation. This feedback loop can be described qualitatively in terms of semi-empirical lumped models. If the jet velocity is such that there is just half a wavelength of the transverse disturbance between the flue and the labium, then the pipe will sound at one of its resonance frequencies (adapted from Fabre and Hirschberg, 2000).

Here, to assess the relevance of such a flute-like mechanism for the production of the bugle's G0 whistling component, we estimate the sounding frequency that would be generated by a jet (of velocity U_j) produced at the level of the supraglottic constriction (of characteristic size H), and hitting a labium consisting of the tensed velum at the level of the intra-pharyngeal ostium located at a distance W from the constriction (see Figs 8, 9). To do this, we calculate a value for the jet velocity U_j based on an estimation of the wapiti's lung capacity, the average duration of bugles (3 s, estimated from our acoustic analyses) and the dimension H (2.5 mm, estimated from dissection). We then estimate the sounding frequency of the whistle using the distance W (70 mm, estimated from our dissections).

Mammalian lung volume V_L increases linearly with increasing body mass M_b (Tenney & Remmers 1963; Worthington et al. 1991; Schmidt-Nielsen 1997) and can be estimated according to the formula:

$$V_L = 0.046 \times M_b^{1.06} \tag{A1}$$

Using this equation, we can estimate the lung volume for a typical male wapiti body mass of 330 kg as 21.5 l. Over 3 s, this corresponds to a typical volume flow of 0.0072 m^3 s^{-1}, forced through an exit channel with a characteristic size H of 2.5 mm, therefore giving a jet velocity U_j of $(0.0072)/(0.0025^2)$=1146 m s^{-1}.

If the perturbation is convected at a velocity of V_j=0.4U_j (Chaigne and Kergomard 2008) along a distance W between the flue exit and the labium estimated at 70 mm, the convection time τ of the acoustic velocity perturbations from the flue exit toward the labium can be estimated as τ=W/V_j=W/(0.4U_j)=1.53×10^{-4} s, corresponding to a characteristic frequency (1/τ) of ~6540 Hz. If the jet velocity is such that there is just half a wavelength of the transverse disturbance between the flue and the labium (first hydrodynamic mode), then the system will sound at half the characteristic frequency, here 3270 Hz.

This sounding frequency has the same order of magnitude as the G0 component of the bugles reported in this study as well as in the literature, confirming that this flute analogy is a plausible hypothesis. If the blowing pressure is increased, shortening the jet travel time, the sounding frequency rises, whereas if the blowing pressure is lowered, the sounding frequency falls. This may explain why towards the end of the bugle, when lung pressure decreases dramatically, the whistle frequency falls towards a minimal G0 value of 160 Hz.

Dimensionless numbers can also be used to assess the compatibility of this system with the production of a flute-like whistle. Reynolds number can be estimated as Re=U_j×H/ν, where ν is the kinematic viscosity of air, which is 1146×0.0025/(1.5×10^{-5}) =1.9×10^5. This value is relatively high compared with most flute-like instruments (10^5 being a typical upper limit). Another relevant dimensionless number, the geometrical ratio W/H, can be estimated as 0.070/0.0025=28, which is again relatively high yet comparable to values characteristic of organ pipes (up to 30; Chaigne and Kergomard, 2008).

Competing interests
The authors declare no competing or financial interests.

Author contributions
D.R. supervised the research. M.T.W. recorded the vocalisations. D.R., D.P. and M.T.W. performed the acoustical and statistical analyses. R.F. and D.R. performed the anatomical and imaging investigations. J.G. and D.R. performed the acoustical modelling. Y.L. provided the specimen and facilitated imaging investigations. D.R., M.T.W., R.F., J.G., D.P. and B.D.C. wrote the manuscript.

Supplementary information
Supplementary information available online at
http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.131219/-/DC1

Data availability
Supporting data are available on request from the corresponding author.

References
Abel, M., Bergweiler, S. and Gerhard-Multhaupt, R. (2006). Synchronization of organ pipes: experimental observations and modeling. *J. Acoust. Soc. Am.* **119**, 2467-2475.

Auregan, Y. and Depollier, C. (1995). Snoring: linear stability analysis and *in vitro* experiments. *J. Sound Vibrat.* **188**, 39-53.

Bachorowski, J.-A. and Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *J. Acoust. Soc. Am.* **106**, 1054-1063.

Boersma, P. and Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.01). Retrieved from http://www.praat.org/.

Bradbury, J. W. (1977). Lek mating behavior in the hammer-headed bat. *Z. Tierpsychol.* **45**, 225-255.

Briefer, E. F., Maigrot, A.-L., Mandel, R., Freymond, S. B., Bachmann, I. and Hillmann, E. (2015). Segregation of information about emotional arousal and valence in horse whinnies. *Sci. Rep.* **5**, 11.

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

Chaigne, A. and Kergomard, J. (2008). *Acoustique des Instruments de Musique.* Paris: Editions Belin.

Charlton, B. D., Ellis, W. A. H., McKinnon, A. J., Cowin, G. J., Brumm, J., Nilsson, K. and Fitch, W. T. (2011). Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: honesty in an exaggerated trait. *J. Exp. Biol.* **214**, 3414-3422.

Charlton, B. D., Frey, R., McKinnon, A. J., Fritsch, G., Fitch, W. T. and Reby, D. (2013). Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Curr. Biol.* **23**, R1035-R1036.

Classe, A. (1957). The whistled language of La Gomera. *Sci. Am.* **196**, 111-120.

Dunn, J. C., Halenar, L. B., Davies, T. G., Cristobal-Azkarate, J., Reby, D., Sykes, D., Dengg, S., Fitch, W. T. and Knapp, L. A. (2015). Evolutionary trade-off between vocal tract and testes dimensions in howler monkeys. *Curr. Biol.* **25**, 2839-2844.

Fabre, B. and Hirschberg, A. (2000). Physical modeling of flue instruments: a review of lumped models. *Acustica* **86**, 599-610.

Fabre, B., Gilbert, J., Hirschberg, A. and Pelorson, X. (2012). Aeroacoustics of musical instruments. *Annu. Rev. Fluid Mech.* **44**, 1-25.

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

Feighny, J. A., Williamson, K. E. and Clarke, J. A. (2006). North American elk bugle vocalizations: male and female bugle call structure and context. *J. Mammal.* **87**, 1072-1077.

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. and 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. and Hauser, M. D. (2002). Unpacking "honesty": generating and extracting information from acoustic signals. In *Animal Communication* (ed. A. Megala-Simmons and A. Popper), pp. 65-137. Berlin: Springer-Verlag.

Fitch, W. T. and Reby, D. (2001). The descended larynx is not uniquely human. *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 1669-1675.

Fletcher, N. H. (2004). A simple frequency-scaling rule for animal communication. *J. Acoust. Soc. Am.* **115**, 2334-2338.

Frey, R. and Gebler, A. (2003). The highly specialized vocal tract of the male Mongolian gazelle (*Procapra gutturosa* Pallas, 1777 - Mammalia, Bovidae). *J. Anat.* **203**, 451-471.

Frey, R. and Riede, T. (2013). The anatomy of vocal divergence in North American elk and European red deer. *J. Morphol.* **274**, 307-319.

Frey, R., Volodin, I. and Volodina, E. (2007). A nose that roars: anatomical specializations and behavioural features of rutting male saiga. *J. Anat.* **211**, 717-736.

Frey, R., Gebler, A., Olson, K. A., Odonkhuu, D., Fritsch, G., Batsaikhan, N. and Stuermer, I. W. (2008). Mobile larynx in Mongolian gazelle: retraction of the larynx during rutting barks in male Mongolian gazelle (*Procapra gutturosa* Pallas, 1777). *J. Morphol.* **269**, 1223-1237.

Frey, R., Volodin, I., Volodina, E., Soldatova, N. V. and Juldaschev, E. T. (2011). Descended and mobile larynx, vocal tract elongation and rutting roars in male goitred gazelles (*Gazella subgutturosa* Güldenstaedt, 1780). *J. Anat.* **218**, 566-585.

Frey, R., Volodin, I., Volodina, E., Carranza, J. and Torres-Porras, J. (2012). Vocal anatomy, tongue protrusion behaviour and the acoustics of rutting roars in free-ranging Iberian red deer stags (*Cervus elaphus hispanicus*). *J. Anat.* **220**, 271-292.

Geist, V. (1999). *Deer of the World: Their Evolution, Behaviour, and Ecology.* UK: Swan Hill Press.

Gerhardt, H. C. (1998). Acoustic signals of animals: recording, field measurements, analysis and description. In *Animal Acoustic Communication: Sound Analysis and Research Methods* (ed. S. L. Hopp, M. J. Owren and C. S. Evans), pp. 1-25. Berlin: Springer-Verlag.

Griffin, D. R. (1958). *Listening in the Dark.* New Haven, CT: Yale University Press.

Griffiths, T. A. (1978). Modification of m. cricothyroideus and the larynx in the Mormoopidae, with reference to amplification of high-frequency pulses. *J. Mammal.* **59**, 724-730.

Harris, T. R., Fitch, W. T., Goldstein, L. M. and Fashing, P. J. (2006). Black and white colobus monkey (*Colobus guereza*) roars as a source of both honest and exaggerated information about body mass. *Ethology* **112**, 911-920.

Hartley, D. J. and Suthers, R. A. (1988). The acoustics of the vocal tract in the horseshoe bat, *Rhinolophus hildebrandti. J. Acoust. Soc. Am.* **84**, 1201-1213.

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

Herbst, C. T., Herzel, H., Svec, J. G., Wyman, M. T. and Fitch, W. T. (2013). Visualization of system dynamics using phasegrams. *J. R. Soc. Interface* **10**, 20130288.

Herzel, H., Berry, D., Titze, I. R. and Saleh, M. (1994). Analysis of vocal disorders with methods from nonlinear dynamics. *J. Speech Hear. Res.* **37**, 1008-1019.

Kelemen, G. and Sade, J. (1960). The vocal organ of the howling monkey (*Alouatta palliata*). *J. Morphol.* **107**, 123-140.

Koda, H., Nishimura, T., Tokuda, I. T., Oyakawa, C., Nihonmatsu, T. and Masataka, N. (2012). Soprano singing in gibbons. *Am. J. Phys. Anthropol.* **149**, 347-355.

McComb, K., Reby, D., Baker, L., Moss, C. and Sayialel, S. (2003). Long-distance communication of acoustic cues to social identity in African elephants. *Anim. Behav.* **65**, 317-329.

McElligott, A. G., Birrer, M. and Vannoni, E. (2006). Retraction of the mobile descended larynx during groaning enables fallow bucks (*Dama dama*) to lower their formant frequencies. *J. Zool.* **270**, 340-345.

Mergell, P., Fitch, W. T. and Herzel, H. (1999). Modeling the role of nonhuman vocal membranes in phonation. *J. Acoust. Soc. Am.* **105**, 2020-2028.

Meyer, J. (2008). Typology and acoustic strategies of whistled languages: phonetic comparison and perceptual cues of whistled vowels. *J. Int. Phonet. Assoc.* **38**, 69-94.

Minami, M. and Kawamichi, T. (1992). Vocal repertoires and classification of the sika deer *Cervus nippon. J. Mammal Soc. Jpn.* **17**, 71-94.

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

Murie, O. J. (1932). Elk calls. *J. Mammal.* **13**, 331-336.

Murie, O. J. (1951). *The Elk of North America.* Harrisburg, PA: Stackpole Books.

Neubauer, J., Edgerton, M. and Herzel, H. (2004). Nonlinear phenomena in contemporary vocal music. *J. Voice* **18**, 1-12.

Nickel, R., Schummer, A. and Seiferle, E. (1987). *Lehrbuch der Anatomie der Haustiere.* Berlin: Paul Parey.

Nowicki, S. and Capranica, R. R. (1986). Bilateral syringeal coupling during phonation of a songbird. *J. Neurosci.* **6**, 3595-3610.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* **41**, 1-16.

Passilongo, D., Reby, D., Carranza, J. and Apollonio, M. (2013). Roaring high and low: composition and possible functions of the Iberian stag's vocal repertoire. *PLoS ONE* **8**, e63841.

R Development Core Team (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Reby, D. and McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim. Behav.* **65**, 519-530.

Rialland, A. (2005). Phonological and phonetic aspects of whistled languages. *Phonology* **22**, 237-271.

Riede, T. and Titze, I. R. (2008). Vocal fold elasticity of the Rocky Mountain elk (*Cervus elaphus nelsoni*): producing high fundamental frequency vocalization with a very long vocal fold. *J. Exp. Biol.* **211**, 2144-2154.

Riede, T., Beckers, G. J. L., Blevins, W. and Suthers, R. A. (2004). Inflation of the esophagus and vocal tract filtering in ring doves. *J. Exp. Biol.* **207**, 4025-4036.

Riede, T., Suthers, R. A., Fletcher, N. H. and Blevins, W. E. (2006). Songbirds tune their vocal tract to the fundamental frequency of their song. *Proc. Natl. Acad. Sci. USA* **103**, 5543-5548.

Sanvito, S., Galimberti, F. and Miller, E. H. (2007). Vocal signalling of male southern elephant seals is honest but imprecise. *Anim. Behav.* **73**, 287-299.

Schmidt-Nielsen, K. (1997). *Animal Physiology: Adaptation and Environment. 5th edn.* Cambridge: Cambridge University Press.

Schön-Ybarra, M. (1995). A comparative approach to the non-human primate vocal tract: implications for sound production. In *Current Topics in Primate Vocal Communication* (ed. E. Zimmermann and J. D. Newman), pp. 185-198. New York: Plenum Press.

Sundberg, J. (1975). Formant technique in a professional female singer. *Acoustica* **32**, 89-96.

Sundberg, J. (1979). Formants and fundamental frequency control in singing: an experimental study of coupling between vocal tract and voice source. *Speech Trans. Lab. Q. Prog. Status Rep.* **20**, 65-78.

Suthers, R. A. (1988). The production of echolocation signals by bats and birds. In *Animal Sonar: Processes and Performance* (ed. P. E. Nachtigall and P. W. B. Moore), pp. 23-45. New York: Plenum Press.

Suthers, R. A. and Fattu, J. M. (1973). Mechanisms of sound production by echolocating bats. *Am. Zool.* **13**, 1215-1226.

Taylor, A. M. and Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *J. Zool.* **280**, 221-236.

Tenney, S. M. and Remmers, J. E. (1963). Comparative quantitative morphology of the mammalian lung: diffusing area. *Nature* **197**, 5456.

Terrien, S., Vergez, C. and Fabre, B. (2013). Flute-like musical instruments: a toy model investigated through numerical continuation. *J. Sound Vibrat.* **332**, 3833-3848.

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

Titze, I. R. and Riede, T. (2010). A cervid vocal fold model suggests greater glottal efficiency in calling at high frequencies. *PLoS Comp. Biol.* **6**, e1000897.

Titze, I. R., Fitch, W. T., Hunter, E. J., Alipour, F., Montequin, D., Armstrong, D. L., McGee, J. and Walsh, E. J. (2010). Vocal power and pressure-flow relationships in excised tiger larynges. *J. Exp. Biol.* **213**, 3866-3873.

Vannoni, E. and McElligott, A. G. (2008). Low frequency groans indicate larger and more dominant fallow deer (*Dama dama*) males. *PLoS ONE* **3**, e3113.

Volodina, E. V., Volodin, I. A., Isaeva, I. V. and Unck, C. (2006). Biphonation may function to enhance individual recognition in the dhole, *Cuon alpinus*. *Ethology* **112**, 815-825.

Weissengruber, G. E., Forstenpointner, G., Peters, G., Kubber-Heiss, A. and 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., Herzel, H., Peters, G. and Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics* **9**, 171-196.

Wiley, R. H. and Richards, D. G. (1978). Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. *Behav. Ecol. Sociobiol.* **3**, 69-94.

Wilson, D. E. and Mittermeier, R. A. (2011). *Handbook of the Mammals of the World*. Vol. **2**. *Hoofed Mammals*. Barcelona: Lynx Edicions.

Worthington, J., Young, I. S. and Altringham, J. D. (1991). The relationship between body mass and ventilation rate in mammals. *J. Exp. Biol.* **161**, 533-536.

Wyman, M. T., Mooring, M. S., McCowan, B., Penedo, M. C. T., Reby, D. and Hart, L. A. (2012). Acoustic cues to size and quality in the vocalizations of male North American bison (*Bison bison*). *Anim. Behav.* **84**, 1381-1391.

Zollinger, S. A., Riede, T. and Suthers, R. A. (2008). Two-voice complexity from a single side of the syrinx in northern mockingbird *Mimus polyglottos* vocalizations. *J. Exp. Biol.* **211**, 1978-1991.