IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Vocal tract shape variation contributes to individual vocal identity in African penguins

This is the author's manuscript

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1944252 since 2025-01-22T11:34:32Z

Published version:

DOI:10.1098/rspb.2023.1029

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

13 September 2026

Vocal tract shape variation contributes to individual vocal identity in African penguins

Livio Favaro[1,2,*,#], Anna Zanoli[2,#], Katrin Ludynia[3,4], Albert Snyman[3], Filippo Carugati[2], Olivier Friard[2], Frine Eleonora Scaglione[5], Luca Manassero[5], Alberto Valazza[5], Nicolas Mathevon[1,6], Marco Gamba[2,*], David Reby[1,6,*]

[1] ENES Bioacoustics Research Lab, CRNL, University of Saint-Etienne, CNRS, Inserm, Saint-Etienne, France

[2] Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

[3] Southern African Foundation for the Conservation of Coastal Birds (SANCCOB), Cape Town, South Africa

[4] Department of Biodiversity & Conservation Biology, University of the Western Cape, Bellville, South Africa

[5] Department of Veterinary Science, University of Turin, Turin, Italy

[6] Institut Universitaire de France, Ministry of Higher Education, Research and Innovation, 1 rue Descartes, 75231 Paris Cedex 05

[#] *These authors contributed equally*

* *Corresponding authors:*

Livio Favaro (livio.favaro@unito.it)

Marco Gamba (marco.gamba@unito.it)

David Reby (david.reby@univ-st-etienne.fr)

ORCID:

Livio Favaro: 0000-0002-8698-472X

Anna Zanoli: 0000-0002-9402-5617

Katrin Ludynia: 0000-0002-0353-1929

Albert Snyman: 0000-0003-4323-2808

Filippo Carugati: 0000-0002-5754-5787

Olivier Friard: 0000-0002-0374-9872

Frine Eleonora Scaglione: 0000-0001-7134-6154

Luca Manassero: 0000-0002-5223-9799

Alberto Valazza: 0000-0002-4745-0653
Nicolas Mathevon: 0000-0003-0219-6601
Marco Gamba: 0000-0001-9545-2242
David Reby: 0000-0001-9261-1711

Abstract

Variation in formant frequencies has been shown to affect social interactions and sexual competition in a range of avian species. Yet, the anatomical bases of this variation are poorly understood. Here, we investigated the morphological correlates of formants production in the vocal apparatus of African penguins. We modelled the geometry of the supra-syringeal vocal tract of 20 specimens to generate a population of virtual vocal tracts with varying dimensions. We then estimated the acoustic response of these virtual vocal tracts and extracted the centre frequency of the first four predicted formants. We demonstrate that (1) variation in length and cross-sectional area of vocal tracts strongly affects the formant pattern, (2) the tracheal region determines most of this variation, and (3) the skeletal size of penguins does not correlate with the trachea length and consequently has relatively little effect on formants. We conclude that in African penguins, while the variation in vocal tract geometry generates variation in resonant frequencies supporting the discrimination of conspecifics, such variation does not provide information on the emitter's body size. Overall, our findings advance our understanding of the role of formant frequencies in bird vocal communication.

Keywords: bioacoustics, formants, source-filter theory, *Spheniscus*, vocal tract modelling

Introduction

Vocal communication is ancient and widespread in vertebrates [1] where it plays a crucial role in speciation, sexual selection [2], and the evolution of complex societies [3]. Numerous studies investigating the function of vocal communication in vertebrates have shown that bird and mammal vocal signals often support individual discrimination or even recognition [4], although in mammals with unstable acoustic cues to identity (e.g., ruminants: [5,6]; felids: [7]; rodents: [8]), updating vocal individuality of social partners can be necessary over time. Individual vocal recognition is also important for developing and mediating social relationships [9-11]. For example, human listeners can recognise familiar individuals by processing the spectral features of their voices [12,13]. More generally, in primates, small differences in the length, volume, and shape of the supra-laryngeal cavities of the vocal tract can lead to small variation in formant patterns that contribute to the encoding of individual identity [14-17]. Such biomechanical and physiological sources of acoustic diversity are broadly shared among mammals [18,19] and have been shown to determine individual spectral signatures in a variety of different species [20,21]. Similarly, formant frequencies have been demonstrated to allow individual recognition in many birds [22,23], although the vocal tract regions involved in filtering the sound source typically differ from those involved in mammalian vocal production [24].

Vocal tract resonances in mammals have also been shown to play a role as cues to body size in the context of sexual selection. While selection pressures appear to favour the emergence of anatomical or behavioural adaptations for size exaggeration, anatomical constraints can ultimately render the vocal signal honest by forcing the length of the resonating cavities to correlate with skeletal size [25]. As a consequence, in mammals, these spectral features have been shown to mediate intersexual competition and facilitate mate choice [26]. In contrast, recent investigations have failed to find similar correlation between formant frequencies and body size [27-29] in birds, suggesting that different anatomical constraints or functional pressures may operate in the Aves class.

When looking at the proximate causes of vertebrate vocal communication, several studies have shown genetic, neural, and anatomical parallels in vocal production between birds and mammals. These include homologies in the brain circuits involved in

the auditory–vocal control [30] and the presence of universal mechanisms for sound production in the bird syrinx and mammalian larynx [31]. However, explorations of the anatomy of bird vocal tracts have revealed a remarkable diversity in morphologies, including elongated and sound-filtering tracheas [32-34] and a complex system of aerial sacs [35]. Such findings suggest that the observed diversification of the bird vocal apparatus has happened since the taxon divergence within the Tetrapod superclass. Comparative studies focussing on the selective forces behind the evolution of the mechanisms involved in generating the format frequencies in birds and mammals might therefore contribute to a broad understanding of the evolution of vertebrate vocal communication.

Penguins are a monophyletic family of seabirds that diverged from the main bird lineage more than 60 Mya [36,37]. Members of this family are philopatric, colonial, and monogamous [38]. Their vocal repertoire is made of four discrete vocal types [39,40], mediating individual recognition and mate choice [41,42]. Penguin calls are produced by vibrating membranes in the syrinx, which generate a complex acoustic signal characterised by a fundamental frequency (corresponding to the rate of vibration of the membranes) and its associated harmonic overtones. The tracheal tube, larynx, and oro-pharyngeal cavity subsequently filter the source signal and generate the formants [43]. Within penguins, the African penguin (*Spheniscus demersus*) is an ideal model species to study the extent to which the resonances of the vocal tract encode meaningful biological information in the avian taxon. Indeed, individuals of this species are highly territorial [39], and because of the selective pressures of the colonial and territorial lifestyle [44], their contact calls and ecstatic display songs (breeding vocalisations) have been found to encode acoustic cues to individual identity in the pattern of the formants [23,24]. Such spectral features of vocalisations are also stable throughout adulthood [45].

Here we investigated how variability in the dimensions of the vocal apparatus of the African penguin may generate cues to identity in the formant patterns of vocalisations. To do this, we collected morphological data from the vocal apparatuses of wild specimens to characterise inter-individual variation and used this data to derive a virtual population of thousands of different vocal tracts. We then modelled the resonances within these computer-generated vocal tracts to identify the regions responsible for generating most of the variability in the acoustic resonances observed

in living penguins. Finally, we investigated the relationship between vocal tract regions and body size, and we discussed whether formant characteristics in vocalisations could also provide reliable information regarding the caller's body size in this species.

Methods

Acoustic recordings and analyses

African penguin vocalisations were recorded in April 2019 in the captive bird exhibit at the Southern African Foundation for the Conservation of Coastal Birds (SANCCOB) in Cape Town, South Africa. Recordings were collected from outside the area at 5-8 m from the vocalising individuals with a Sennheiser MKH 416 P48 directional microphone (frequency response 20 Hz to 20 kHz) covered with a Rycote Classic Softie Windshield and connected to a ZOOM H5 handy recorder (48 kHz sampling rate). Audio files were saved in RIFF-WAVE format (16-bit amplitude resolution) and stored on a 32GB secure digital (SD) memory card.

Visual examination of the spectrograms allowed the identification of 95 ecstatic display songs (i.e., loud sequences combining three types of syllables [40]) showing a good signal-to-noise ratio and emitted by 12 adult individuals (11 males and one female) to mediate mate choice and territorial defence. The individuals were sexed by genetic analyses and then identified with flipper bands already provided by SANCCOB. The contribution of each penguin to the final dataset is provided in the Supplementary Material (Table S1).

The contour of the first four formants (F_1– F_4) of the first type-B syllable (i.e., the longest vocal units within a sequence [46]) of each display song was extracted in Praat (V. 6.1.40) [47] using a Linear Predictive Coding analysis (maximum number of formants = 5, maximum formant = 4000 Hz, time step = 0.045 s). The accurate tracking of the formants was ensured by comparing the output of the LPC with a visual examination of the spectrograms and correcting for octave jumps when necessary. For each of the first four formants (F_1– F_4), we calculated the interindividual coefficient of variation (CV = ${}^{F_n\ \mathrm{SD}}/_{F_n\ \mathrm{mean}}$).

Specimens and casting of the vocal tracts

We investigated the vocal apparatus of 20 adult African penguins (4 males and 16 females) found dead in April and May 2019 along the coasts of the South African province of the Western Cape. Although there was no precise way to accurately

158 determine the date of death of individuals, at the time of collection, all cadavers were
159 in a good state of preservation with no signs of scavenging or secondary predation
160 present. All post-mortems examinations were conducted at the SANCCOB within 48
161 hours after being collected from the wild, ensuring that all samples were in the best
162 state possible. Cadavers were sexed by visual inspection of the gonads (ovaries or
163 testes). However, in the African penguin, no sexual dimorphism has been documented
164 in vocal anatomy that could affect EDS's spectral characteristics. Thus, we did not use
165 the sex of individuals as a variable in subsequent analyses. We obtained a silicone cast
166 of the vocal apparatus of each penguin by injecting high-speed catalysed silicone rubber
167 (Bluestar Silicones Rhodorsil® RTV) through the larynx while closing, with surgical
168 clamps, the bronchial tubes shortly after the syrinx. Once the entire supra-syringeal
169 vocal tract was filled with the silicone solution, we filled in the oral region, placing the
170 bird in the natural position observed during the emission of the ecstatic display songs.
171 This procedure gave us a reliable cast of the entire vocal apparatus during phonation
172 [48] that we extracted after solidification in that precise position (approximately 5 hrs
173 later) as a single piece through the mouth (Figure 1a). In the silicon cast the
174 laryngopharyngeal region is not visible because in cadavers the laryngeal opening is
175 closed as it has collapsed (Suppl. Figure S1). To use it in the models and in all
176 computational aspects of the work, we manually placed the structures in the natural
177 phonatory position and extracted the relevant measurements with a calliper. Finally, for
178 each bird, we measured the inter-skull length (ISL) as a proxy of the skeletal dimension
179 [49,50]. All the measurements were taken by the same person (A.S.) using a digital
180 calliper accurate to 0.005 mm.
181
182 *Computed tomography scan and calculation of the cross-sectional area*
183 Silicone casts of the vocal apparatuses were shipped to the Department of
184 Veterinary Science of the University of Turin (Italy) for Computed Tomography (CT)
185 imaging. All CT images were acquired in a 16-slice helical scanner (SOMATOM Emotion
186 16) with slices of 2 mm thickness. Later, the DICOM files were visualised and post-
187 processed as multi-planar reconstruction (MPR) with the processing software Osirix
188 (version 6.0, 64bits for Mac, www.osirix-viewer.com). Using the '3D curved path'
189 feature on the midsagittal plane, we measured the cross-sectional area of each vocal
190 tract at 2 mm slices along its length from above syringeal constriction until the beak

opening (Figure 1). We did not include the beak length in the length of the vocal tract because the EDS is produced with a fully opened beak [40].

Computational modelling of the vocal tracts

We used a multi-tube model to generate the vocal tract models that simulated the vocal tract as a series of cylindrical tubes concatenated end-to-end [51]. For all calculations, we considered a single tracheal pipe from the syrinx until the double trachea septum ended into a uniform tracheal region [52]. Once we identified portions' cross-sectional areas for each of the penguins, we estimated formants for each specimen using the anatomical measurements of its vocal tract, and we calculated inter-individual coefficient of variation (CV) for each of the first four formants (F_1– F_4). Furthermore, in order to estimate the fit of the formant patterns predicted by our vocal tract models, we performed a regression analysis with the formants estimated from the vocal tract geometries as a predictor variable and the formants measured in the actual EDS recordings as the predicted variable. The slope of this regression gives a fit for the scaling (the "formant density" as determined by vocal tract length), and the Pearson's coefficient provides a fit for the pattern (the relative position of the formants independent of vocal tract length).

Having verified that the vocal tract models yielded results within the range of variation of natural formants, we automatically increased or decreased lengths and areas of those portions belonging to a particular region of the vocal tract. We did these changes within the overall variation of that specific region, as measured from the specimens (5% to 10% for the single tracheal pipe, 5% to 20% for the trachea, 5% to 20% for the laryngopharyngeal region, and 5% to 15% for the oral cavity). We generated a virtual population of 1160 penguin vocal tracts using these vocal tract area functions as input for a customised version of VTAR modelling software [53,54]. We calculated the acoustic response of each model and collected formants F_1, F_2, F_3, and F_4, total length, and total volume, of the region whose measurements were changed [55,56].

Statistical analyses

Following the approach proposed by Reby and co-authors [19], we confirmed the reliability of the computational models in predicting the resonances of the penguins'

vocal tracts by fitting a linear regression between the resonances of the vocal tracts observed in the ecstatic display songs of living African penguins and those predicted for the original silicone casts subjected to CT scan.

Subsequently, to investigate whether region variation was responsible for generating most of the variability in the formant model, we ran eight separate Linear Mixed Models (LMMs [57]) in R (version 4.1.2 [58]). We considered the effect of regions' length and area variation separately, keeping the natural cross-sectional areas for the former (model 1a, 2a, 3a and 4a) and the natural lengths for the latter (model 1b, 2b, 3b and 4b). In each group of models, the response variables were the logarithmic transformation of F_1 (models 1a and 1b), F_2 (models 2a and 2b), F_3 (models 3a and 3b) and F_4 (models 4a and 4b). We scaled the length and the area variation for each region, the total length, and volume, thus creating z-scored variables. We built the models using the z-scored total length and the interaction between the z-scored length variation and the relative region as fixed factors when considering only the length variation. Likewise, the interaction between the z-scored area variation and the relative region were the fixed factors of the models considering only the area variation. We used the z-scored total volume as a control predictor for all the models and the penguin identity as a random factor.

We applied the logarithmic transformation to the vocal tract length and the formant values to obtain a normal distribution and homogeneity of the residuals after looking at the qq-plot and plotting the residuals against the fitted values [59]. We estimated multicollinearity with generalised variance inflation factors (GVIF [60]; R-function 'vif' [61]). We excluded the collinearity between fixed factors and any parameter estimation issue (all VIFs < 4.8 [62]). We tested the models' significance by comparing each full model against a control model, including only the control predictor and the random factor [63], using a likelihood ratio test (Anova with the 'Chisq' test argument [64]). Then, through likelihood ratio tests between the full model and the respective control model, we calculated p-values for each predictor (R-function' drop1' [65]). We performed post hoc pairwise comparisons using the 'lstrends' and the 'pairs' functions ('lsmeans' R-package [66]) to identify whether the regions' length and area variation produced a significant effect on the formant pattern.

Finally, to investigate whether the regions responsible for generating most of the variability in the formant pattern were unconstrained by the skeletal dimension, we performed Pearson's correlation test ('cor.test' R-function) between the inter-skull length (mm) and each region's length measured from the original silicon cast of the

cadavers' vocal tracts. Then, we also tested the correlation between the inter-skull length (mm) and the original F_n values measured from the silicon cast of the cadavers' vocal tracts.

Results

Formants in actual penguins' vocalisations vs predicted from modelled vocal tracts

Our results confirmed that the formants observed in living birds fitted well with those obtained by simulating the acoustic resonances in the modelled vocal tracts. Indeed, both the slope of the regression line (indicating a better fit of the scaling of the resonances) and R^2 (indicating a better fit of the pattern of the resonances) were close to 1 (Figure 2). In particular, the formants values measured for the ecstatic display songs were, on average, F_1=812 Hz (CV=8.67%), F_2=1309 Hz (CV=4.33%), F_3=1851 Hz (CV=4.52%), and F_4=2733 Hz (CV=5.99%), while formants predicted from the vocal tract geometries were, on average, F_1=617 Hz (CV=4.71%), F_2=1390 Hz (CV=5.96%), F_3=2192 Hz (CV=6.46%), and F_4=2908 Hz (CV=9.01%), showing an average variation of less than 10% in frequency.

Vocal tract length variation

The models we built to investigate whether, using the natural cross-sectional areas, the length variation of the vocal tract regions influenced the formant pattern significantly differed from the control models (Table 1). We found a negative effect of the vocal tract total length on formant frequencies (Figure 3a). By examining the impact of each region length variation on the formants, we found that the elongation of the single tracheal pipe region lowered F_1 and F_2 and raised F_3 but did not impact F_4. The lengthening of the laryngopharyngeal region and of the trachea decreased the formants. The elongation of the oral cavity increased only F_2, F_3, and F_4 (Figure 3c; Table S2a). The pairwise comparison between the length variation of the regions revealed that, across all regions, the elongation of the single tracheal pipe produced the most substantial change in formant frequencies. To a minor extent, lengthening the laryngopharyngeal region also led to a more pronounced change of formants than the trachea and the oral cavity (Figure 3c). We reported estimates, standard error, and p-values for all the comparisons in Table S3a.

Vocal tract cross-sectional area variation

The models we built to investigate whether, using the natural lengths, the area variation of the vocal tract regions influenced the formant pattern significantly differed from the control models (Table 2). By testing the effect of the variation of each region area on the formant frequencies, we found that a widening of the single tracheal pipe region lowered F_1, F_2, and F_3 but raised F_4. Widening the laryngopharyngeal region increased all the formants, while widening the trachea increased F_1 and F_2 and decreased F_3 and F_4. Widening the oral cavity area increased only F_3 and F_4 (Figure 3c; Table S2b). The pairwise comparison between the area variation of the regions revealed that, across all regions, the widening of the single tracheal pipe produced the most substantial change in formant frequencies, for the exclusion of F_4. To a minor degree, widening the laryngopharyngeal region also led to a more pronounced change of formants than the trachea and the oral cavity (Figure 3c). We reported estimates, standard error, and p-values for all the comparisons in Table S3b.

Vocal tract length, formants, and body size

When testing the correlation between the inter-skull length (mm) and the regions' length, we found that this proxy of body size correlated only with the length of the oral cavity (Pearson's correlation: cor = 0.55, t = 2.767, df = 18, p-value = 0.012). For all the other regions, we did not find a significant correlation with the inter-skull length (ISL vs. single tracheal pipe: t = 1.814, df = 18, p-value = 0.08; vs. trachea: t = 1.029, df = 18, p-value = 0.31; vs. laryngopharyngeal region: t = -0.303, df = 18, p-value = 0.76). In addition, there were no significant correlations between the inter-skull length and any formant (ISL vs. F_1: t = -1.445, df = 18, p-value = 0.16; vs. F_2: t = -1.655, df = 18, p-value = 0.11; vs. F_3: t = -1.690, df = 18, p-value = 0.10; vs. F_4: t = -1.293, df = 18, p-value = 0.21).

Table 1. Effect of the length variation of the vocal tract regions on the formant pattern F_1-F_4, using the natural cross-sectional areas (NA). Full vs. Control: Model 1a: Chisq = 3180.721, df = 8, p-value < 0.0001; Model 2a: Chisq = 2847.599, df = 8, p-value < 0.0001; Model 3a: Chisq = 1889.097, df = 8, p-value < 0.0001; Model 4a: Chisq = 1230.825, df = 8, p-value < 0.0001. [a] Not shown as having a meaningful interpretation. [b] Estimate ± SE refers to the difference in response between the reported level of this categorical predictor and the reference category of the same predictor. [c] These predictors were dummy coded, with the 'single tracheal pipe x z length variation' being the reference category. "-" denotes a missing value.

		Model 1a			Model 2a			Model 3a			Model 4a		
	response variables	**log10(F_1)**			**log10(F_2)**			**log10(F_3)**			**log10(F_4)**		
	fixed factors	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**
NA	**z total length**	-0.011	0.001	<0.0001	-0.017	0.001	<0.0001	-0.030	0.002	<0.0001	-0.023	0.003	<0.0001
	z total volume	0.003	0.000	a	0.004	0.001	a	0.000	0.001	a	-0.027	0.002	a
	laryngopharyngeal region	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a
	oral cavity	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a
	trachea	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a	-0.001	0.000	a
	z length variation	-0.036	0.001	a	-0.022	0.002	a	0.007	0.003	a	0.006	0.005	a
	region x z length variation	-	-	<0.0001	-	-	<0.0001	-	-	<0.0001	-	-	<0.0001
	laryngopharyngeal region x z length variation[b,c]	0.035	0.001	-	0.021	0.002	-	-0.009	0.003	-	-0.008	0.005	-
	oral cavity x z length variation[b,c]	0.036	0.001	-	0.022	0.002	-	-0.005	0.003	-	0.000	0.005	-
	trachea x z length variation[b,c]	0.036	0.001	-	0.022	0.002	-	-0.008	0.003	-	-0.007	0.005	-

Table 2. Effect of the vocal tract region's area variation on the formant pattern F_1-F_4, using the natural lengths (NL). Full vs. Control: Model 1b: Chisq = 1230.268, df = 7, p-value < 0.0001; Model 2b: Chisq = 1364.831, df = 7, p-value < 0.0001; Model 3b: Chisq = 1002.692, df = 7, p-value < 0.0001; Model 4b: Chisq = 488.873, df = 7, p-value < 0.0001.) [a] Not shown as having a meaningful interpretation. [b] Estimate ± SE refers to the difference in response between the reported level of this categorical predictor and the reference category of the same predictor. [c] These predictors were dummy coded, with the 'single tracheal pipe x z area variation' being the reference category. "-" denotes a missing value.

		Model 1b			Model 2b			Model 3b			Model 4b		
	response variables	**log10(F_1)**			**log10(F_2)**			**log10(F_3)**			**log10(F_4)**		
	fixed factors	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**	**estimate**	**SE**	**p-value**
NL	**z total volume**	0.002	0.001	a	0.002	0.001	a	0.001	0.001	a	-0.010	0.001	a
	laryngopharyngeal region	0.000	0.000	a	0.000	0.000	a	-0.001	0.000	a	0.000	0.000	a
	oral cavity	0.000	0.000	a	0.000	0.000	a	0.000	0.000	a	0.000	0.000	a
	trachea	0.000	0.000	a	0.000	0.000	a	0.000	0.000	a	0.000	0.000	a
	z area variation	-0.004	0.000	a	-0.005	0.000	a	-0.003	0.000	a	0.003	0.000	a
	region x z area variation	-	-	<0.0001	-	-	<0.0001	-	-	<0.0001	-	-	<0.0001
	laryngopharyngeal region x z area variation[b,c]	0.006	0.000	-	0.007	0.000	-	0.006	0.000	-	-0.001	0.000	-
	oral cavity x z area variation[b,c]	0.004	0.000	-	0.005	0.000	-	0.004	0.000	-	0.001	0.000	-
	trachea x z area variation[b,c]	0.005	0.000	-	0.005	0.000	-	0.002	0.000	-	-0.005	0.000	-

Discussion

Using an integrated approach combining anatomical investigations, vocal tract modelling and acoustic analyses, we examined how variation in the geometry (length and cross-sectional areas) of the vocal tract generates variability in formant patterns of the EDS vocalisations of African penguins. Our results provide the first evidence in a seabird species that 1) the variation in the length and cross-sectional area of the vocal tract affects formant frequency patterns and 2) this variation is better described considering the independent contribution of the different vocal tract regions, rather than the overall volume of the vocal tract. When looking at the models simulating random variation in the morphology of the penguins' vocal tracts, we found that an increase in the length of the whole tracheal tube (i.e., single tracheal pipe + trachea + laryngopharyngeal region) lowers most of the formant frequencies. This is consistent with the results of a study modelling the acoustic resonances in the vocal tract of the house sparrow (*Passer domesticus*), where decreasing the tracheal length was found to increase the harmonic resonance [67]. However, in the African penguin, when considering the length variation of specific regions, we found that the elongation of the oral cavity increased formant frequencies, although the effect is minimal (TABLE S2a). Moreover, we found that widening the single tracheal pipe in the African penguin lowers the formants, but the same variation applied to the other regions determines an overall increase in the resonant frequencies. Also in this case, our findings support what has been described in songbirds' vocal system, where the oropharyngeal-oesophageal cavity strongly influences sound filtering [67-69].

Formant frequencies have been shown to provide acoustic cues to caller identity in several non-passerine birds [23,24,27-29]. However, the anatomical determinants of the formant patterns have never been investigated in detail. Here, we report that, in the African penguin, slight variations (i.e., 5% to 20%) of the length and cross-sectional area of the tracheal regions - either in the portion where it is divided by the longitudinal septum [52] or in the proximity of the laryngopharyngeal cavity - cause a broad shift (i.e., up to 40%; Figure 3) in the formants of the vocalisations. We thus conclude that these regions can play a crucial role in determining the formant pattern of penguins' calls and, thus, in encoding the acoustic cues to individual identity in this species [24]. Previous studies demonstrated that humans perceive shifts in formant frequency spacing in speech of 4% or less [70,71], while non-human mammals are sensitive to

shifts of 8–10% variation in their species-specific calls [72-74]. Although further investigations are needed to precisely determine most bird species' sensitivity to variation in formant frequencies, experimental evidence suggests that a variation of ±20% of the formant spacing allows African penguins to perceive variation between conspecific vocalisations [75]. This species lives in relatively dense colonies, but each pair has a defined nest and only interacts with direct neighbours. For this reason, a small Potential of Individual Coding (PIC: the ratio between the within-individual variation and the between-individual variation) is necessary [24], unlike species that live in large colonies but do not nest [39].

Importantly, we also found that the lengths of most of the regions of the vocal tract do not correlate with the birds' skeletal dimension (i.e., inter-skull length - ISL) and that ISL does not correlate with formants predicted for their vocal tracts. The only portion of the vocal tract that was significantly correlated with inter-skull length was the oral cavity, which only contributed to ~20% of the formant variability of the spectral envelope (Figure 3c). Together, our results show that, in the African penguin, the anatomical regions of the vocal tract affecting formant correlates of individual identity are not determined by the body dimension of the caller. Formant frequencies are known to provide cues to caller body size in several non-human mammals, including ungulates [25], non-human primates [76,77], and even marsupials [78]. In these taxa, the length of the vocal tract is constrained to various extents by the skeletal size, including the neck, head, and sternum. However, in cases where the vocal tract development is relatively free from skeletal size constraints, such as in humans, where the larynx is descended from the normal mammal position as an adaptation for speech, only a weak relationship is observed between formant frequencies and body dimension when examining individuals from the same sex and age class [77,79,80]. Similarly, we found that a lack of correlation between format frequencies and skeletal size is observed in the African penguin, where the bird's skeleton does not anatomically constrain the trachea. Indeed, while it has been found that in the King penguin (Genus *Aptenodytes*) the trachea lies straight in line with the thorax [81] in the African penguin, the trachea and oesophagus form a slight loop as they exit the interclavicular air sac and are laterally displaced towards the right-hand side of the neck (Suppl. Figure S2). To the best of our knowledge, the lack of correlation between formant frequencies and skeletal dimension has been reported for four other bird species, including the Humboldt

(*Spheniscus humboldti*) and Magellanic (*S. magellanicus*) penguins [29], two congeneric species of the African penguin, the herring gull (*Larus argentatus* [28]), and the phylogenetically distant corncrake (*Crex crex* [27]). Our findings, combined with this previous research, suggest that this anatomical condition in birds may be ancestral and widespread also beyond species with "exaggerated" trachea elongation [33]. Overall, our results also broaden our knowledge of the role of *homologous* anatomical regions of bird and mammal vocal apparatuses in shaping the spectral envelope of the vocalisations [82,83].

More generally, our results indicate that, in the African penguin, the filter-related information encoded in the spectral envelope of the callers' vocalisations is likely to support the ability of receivers to discriminate conspecifics and recognize their reproductive partner [24]. However, the lack of clear correlation between body size indexes and formant frequencies may explain why the animals do not appear to use formant frequency scaling to assess the size of potential mates or rivals. Further investigations should be directed toward understanding whether the source-related vocal parameters (determined by the dimension of the syringeal membranes) or the temporal characteristics of the vocal displays (determined by the lungs and air sacs capacity) might instead convey honest information about the emitter's body size and thus mediate intrasexual competition and intersexual mate choice.

Ethics. The study was approved by the Research and Ethics Committee of The South African Foundation for the Conservation of Coastal Birds (approval number REC19/02).

Data Accessibility. The data are available upon request at https://doi.org/10.5281/zenodo.7867240

Authors' contributions. LF, MG, NM, and DR conceived and designed the study. LF realised the silicone casts with the help of AS, KL and DR. AV, LM, and FS performed the computed tomography scans with the help of LF and MG. FC collected the morphological measurements in Osirix. OF and MG realised the computational models of the vocal tracts. AZ and MG performed the statistical analyses. AZ, LF, MG, and DR wrote the paper with helpful input from all authors.

Competing interests. We declare we have no competing interests.

Funding. This study was funded by the University of Lyon IDEXLYON [grant no. ANR-16-IDEX-0005], Centre National de la Recherche Scientifique (CNRS), University of Lyon/Saint-Etienne, Institut Universitaire de France (NM & DR), and Labex CeLyA.

SANCCOB is supported by a large variety of funders. This research was also partially supported by the National Geographic Society [grant number NGS-67401R-20].

Acknowledgements. We are grateful to Eleonora Cresta for her support during the silicone casting procedures and to Federica Sammartano for helping with computed tomography scans. We also thank Franziska Feist and Emma Tourtigues for their contribution to the visual inspection of the spectrograms.

Figure Legends

Figure 1. Silicon cast of an African penguin vocal tract. Lines crossing the cast indicate the position of three slices of CT scans represented in the panels: (a) oral cavity, (b) trachea, (c) tracheal pipes. The trachea is divided into two pipes starting just below slice b (i.e., approx. 1 cm below the larynx). The laryngopharyngeal region is not visible in the silicon cast because in cadavers the laryngeal opening is closed.

Figure 2. Correlations between the resonances measured in African penguin ecstatic display songs (y-axis) and those predicted for the CT scanned vocal tracts (x-axis).

Figure 3. (a) Effect of the variation of the vocal tract regions on formant frequencies when using natural cross-sectional area measurements. (b) Impact of the vocal tract region volume on formant frequencies when using natural length measurements. (c) Percentage of F_n variation [$-25<\Delta F_n(\%)<+25$] when lengthening and widening each region of the vocal tract. Purple circles represent the variation range of ΔF_n (minimum ΔF_n - maximum ΔF_n) under the region length variation condition; Blue circles represent the variation range of ΔF_n (minimum ΔF_n - maximum ΔF_n) under the region area variation condition.

References

1. Chen Z, Wiens JJ. 2020 The origins of acoustic communication in vertebrates. *Nat. Commun.* 11,369. https://doi.org/10.1038/s41467-020-14356-3
2. Wilkins MR, Seddon NR, Safran RJ. 2013 Evolutionary divergence in acoustic signals: causes and consequences. *Trends. Ecol. Evol.* 28,156–166. https://doi.org/10.1016/j.tree.2012.10.002

3. Fitch WT. 2017 Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* 24,3–33. https://doi.org/10.3758/s13423-017-1236-5
4. Carlson NV, Kelly EMK, Couzin I. 2020 Individual vocal recognition across taxa: a review of the literature and a look into the future. *Phil. Trans. R. Soc. B.* 375(1802),20190479. https://doi.org/10.1098/rstb.2019.0479
5. Briefer E, Vannoni E, McElligott AG. 2010 Quality prevails over identity in the sexually selected vocalisations of an ageing mammal. *BMC Biol*. 8, 35. https://doi.org/10.1186/1741-7007-8-35
6. Sibiryakova OV, Volodin IA, Matrosova VA, Volodina EV, Garcia AJ, Gallego L, Landete-Castillejos T. 2015 The power of oral and nasal calls to discriminate individual mothers and offspring in red deer, Cervus elaphus. *Front. Zool.* 12, 2. https://doi.org/10.1186/s12983-014-0094-5
7. Chelysheva EV, Klenova AV, Volodin IA, Volodina EV. 2023 Advertising sex and individual identity by long-distance chirps in wild-living mature cheetahs (*Acinonyx jubatus*). *Ethology*. 129, 288-300. https://doi.org/10.1111/eth.13366
8. Matrosova VA, Volodin IA, Volodina EV. 2009 The short-term and long-term individuality in speckled ground squirrel alarm calls. *J. Mammal.* 90, 158-166. https://doi.org/10.1644/08-MAMM-A-032.1
9. Rendall D, Cheney DL, and Seyfarth RM. 2000 Proximate factors mediating "contact" calls in adult female baboons (*Papio cynocephalus ursinus*) and their infants. *J. Comp. Psychol. 114*(1),36. https://doi.org/10.1037/0735-7036.114.1.36
10. McComb K, Moss C, Sayialel S, Baker L. 2000 Unusually extensive networks of vocal recognition in African elephants. *Anim. Behav.* 59(6),1103-1109. https://doi.org/10.1006/anbe.2000.1406
11. Ward A, Webster M. 2016 Mechanisms: Social Recognition and Social Organisation. In: *Sociality: The Behaviour of Group-Living Animals.* Springer, Cham. https://doi.org/10.1007/978-3-319-28585-6_2
12. Lavan N, McGettigan C. 2023 A model for person perception from familiar and unfamiliar voices. *Commun. Psychol.* 1, 1. https://doi.org/10.1038/s44271-023-00001-4
13. Scott S, McGettigan C. 2016 The voice: From identity to interactions. In *APA handbook of nonverbal communication* (eds D Matsumoto, HC Hwang, MG Frank) pp. 289–305. American Psychological Association. https://doi.org/10.1037/14669-011
14. Vorperian HK, Wang S, Chung MK, Schimek EM, Durtschi RB, Kent RD, ... Gentry LR. 2009 Anatomic development of the oral and pharyngeal portions of the vocal

tract: An imaging study. *JASA. 125*(3), 1666-1678. https://doi.org/10.1121/1.3075589

15. Mathias SR, von Kriegstein K. 2014 How do we recognise who is speaking? *Front. Biosci.* 6, 92–109. https://doi.org/10.2741/s417
16. Gamba M, Favaro L, Araldi A, Matteucci V, Giacoma C, Friard O. 2017 Modeling individual vocal differences in group-living lemurs using vocal tract morphology. *Curr. Zool.* 63(4), 467-475. https://doi.org/10.1093/cz/zox023
17. Wyman MT, Walkenhorst B, Manser MB. 2022. Selection levels on vocal individuality: strategic use or byproduct. *Curr. Opin. Behav. Sci.* 46, 101140. https://doi.org/10.1016/j.cobeha.2022.101140
18. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016 Monkey vocal tracts are speech-ready. *Sci. Adv.* 2, e1600723-e1600723. https://doi.org/10.1126/sciadv.1600723
19. Reby D, Wyman MT, Frey R, Charlton BD, Dalmont JP, Gilbert J. 2018 Vocal tract modelling in fallow deer: are male groans nasalized?. *J. Exp. Biol.* 221(17), jeb179416. https://doi.org/10.1242/jeb.179416
20. Townsend SW, Charlton BD, Manser MB. 2014 Acoustic cues to identity and predator context in meerkat barks. *Anim. Behav.* 94, 143-149. https://doi.org/10.1016/j.anbehav.2014.05.021
21. Green A, Clark C, Favaro L, Lomax S, Reby D. 2019 Vocal individuality of Holstein-Friesian cattle is maintained across putatively positive and negative farming contexts. *Sci. Rep.* 9(1), 1-9. https://doi.org/10.1038/s41598-019-54968-4
22. Fitch WT, Kelley JP. 2000 Perception of Vocal Tract Resonances by Whooping Cranes Grus americana. *Ethology*. 106(6), 559-574. https://doi.org/10.1046/j.1439-0310.2000.00572.x
23. Favaro L, Gili C, Da Rugna C, Gnone G, Fissore C, Sanchez D, ... Pessani D. 2016 Vocal individuality and species divergence in the contact calls of banded penguins. *Behav Processes*. 128,83-88. https://doi.org/10.1016/j.beproc.2016.04.010
24. Favaro L, Gamba M, Alfieri C, Pessani D, McElligott AG. 2015 Vocal individuality cues in the African penguin (*Spheniscus demersus*): a source-filter theory approach. *Sci. Rep.* 5(1),17255. https://doi.org/10.1038/srep17255
25. Reby D, 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. https://doi.org/10.1006/anbe.2003.2078

26. Reby D, McComb K, Cargnelutti B, Darwin C, Fitch WT, Clutton-Brock T. 2005 Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proc. Roy. Soc. B.* 272(1566), 941-947. https://doi.org/10.1098%2Frspb.2004.2954
27. Budka M, Osiejuk TS. 2013 Formant frequencies are acoustic cues to caller discrimination and are a weak indicator of the body size of corncrake males. *Ethology*. 119, 960–969. https://doi.org/10.1111/eth.12141
28. Hardouin LA, Thompson R, Stenning M, Reby D. 2014 Anatomical bases of sex- and size-related acoustic variation in herring gull alarm calls. *J. Avian. Biol.* 45, 157-166. https://doi.org/10.1111/j.1600-048X.2013.00144.x
29. Favaro L, Gamba M, Gili C, Pessani D. 2017 Acoustic correlates of body size and individual identity in banded penguins. *PLoS One*. 12(2), e0170001. https://doi.org/10.1371/journal.pone.0170001
30. Jarvis ED. 2007 Neural Systems For vocal learning in birds and humans: a synopsis. *J. Ornithol.* 148, 35–54. https://doi.org/10.1007/s10336-007-0243-0
31. Elemans CP, Rasmussen JH, Herbst CT, Düring DN, Zollinger SA, Brumm H, ... Švec JG. 2015 Universal mechanisms of sound production and control in birds and mammals. *Nat. Commun. 6*(1), 8978. https://doi.org/10.1038/ncomms9978
32. Beckers GJ, Suthers RA, Cate CT. 2003 Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc. Natl. Acad. Sci. U.S.A.* 100(12), 7372-7376. https://doi.org/10.1073%2Fpnas.1232227100
33. Fitch WT. 1999 Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *J. Zool.* 248, 31–48. https://doi.org/10.1111/j.1469-7998.1999.tb01020.x
34. Daley M, Goller F. 2004 Tracheal length changes during zebra finch song and their possible role in upper vocal tract filtering. *J. Neurobiol.* 59, 319–330. https://doi.org/10.1002/neu.10332
35. Schmidt MF, Wild JM. 2014 The respiratory-vocal system of songbirds: anatomy, physiology, and neural control. *Prog. Brain. Res.* 212, 297-335. https://doi.org/10.1016%2FB978-0-444-63488-7.00015-X
36. Baker AJ, Pereira SL, Haddrath OP, Edge KA. 2006 Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proc. R. Soc. B.* 273, 11–17. https://doi.org/10.1098%2Frspb.2005.3260

570 37. Vianna JA, Fernandes FA, Frugone MJ, Figueiró HV, Pertierra LR, Noll D, ... Bowie
571 RC. 2020 Genome-wide analyses reveal drivers of penguin diversification. *Proc. Natl.*
572 *Acad. Sci*. 117(36), 22303-22310. https://doi.org/10.1073/pnas.2006659117
573 38. Borboroglu G, Boersma P. 2013 Penguins: natural history and conservation. Seattle:
574 University of Washington Press.
575 39. Jouventin P. 1982 Visual and vocal signals in penguins, their evolution and adaptive
576 characters. *Fortschritte der Verhaltensforschung*. *24,* 148.
577 40. Favaro L, Ozella L, Pessani D. 2014 The vocal repertoire of the African Penguin
578 (*Spheniscus demersus*): structure and function of calls. *PloS One*. *9*(7), e103460.
579 https://doi.org/10.1371/journal.pone.0103460
580 41. Aubin T, Jouventin P, Hildebrand C. 2000 Penguins use the two–voice system to
581 recognize each other. *Proc. R. Soc. B.* 267:1081–1087.
582 https://doi.org/10.1098/rspb.2000.1112
583 42. Baciadonna L, Solvi C, La Cava S, Pilenga C, Gamba M, Favaro L. 2021 Cross-modal
584 individual recognition in the African penguin and the effect of partnership. *Proc. R.*
585 *Soc. B.* 288(1960), 20211463. https://doi.org/10.1098/rspb.2021.1463
586 43. Favaro L, Gamba M, Ludynia K, Snyman A, Pichegru L, Cresta E, ... Reby D. 2020,
587 December. Computational modelling of penguins' vocal tract. In: *Forum*
588 *Acusticum* (pp. 2037-2037).
589 44. Jouventin P, Aubin T. 2002 Acoustic systems are adapted to breeding ecologies:
590 individual recognition in nesting penguins. *Anim. Behav.* 64(5), 747-757.
591 https://doi.org/10.1006/anbe.2002.4002
592 45. Calcari C, Pilenga C, Baciadonna L, Gamba M, Favaro L. 2021 Long-term stability of
593 vocal individuality cues in a territorial and monogamous seabird. *Anim. Cogn.* 24(6),
594 1165-1169. https://doi.org/10.1007/s10071-021-01518-z
595 46. Favaro L, Gamba M, Cresta E, Fumagalli E, Bandoli F, Pilenga C, ... Reby D. 2020
596 Do penguins' vocal sequences conform to linguistic laws?. *Biol. Lett.* 16(2),
597 20190589. https://doi.org/10.1098/rsbl.2019.0589
598 47. Boersma P. 2001 Praat, a system for doing phonetics by computer. *Glot Int.* 5, 341–
599 345.
600 48. Favaro L, Pichegru L. 2018 Penguins: Behavioural Ecology and Vocal
601 Communication. In *Encyclopedia of Animal Cognition and Behavior* (eds J Vonk, T
602 Shackelford) Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_844-1

603 49. Plateau O, Foth C. 2021 Common Patterns of Skull Bone Fusion and Their Potential
604 to Discriminate Different Ontogenetic Stages in Extant Birds. Front. Ecol. Evol. 9,
605 737199. https://doi.org/10.3389/fevo.2021.737199
606 50. Shatkovska OV, Ghazali M. 2021 Relative skull size as one of the factors limiting
607 skull shape variation in passerines. Can. J. Zool. 99(12), 1054-1066.
608 https://doi.org/10.1139/cjz-2021-0106
609 51. Gamba M, Friard O, Giacoma C. 2012 Vocal tract morphology determines species-
610 specific features in vocal signals of lemurs (*Eulemur*). *Int. J. Primatol.* 33, 1453-
611 1466. https://doi.org/10.1007/s10764-012-9635-y
612 52. Zeek PM. 1951 Double trachea in penguins and sea lions. *Anat. Rec.* 111(3),327-
613 43. https://doi.org/10.1002/ar.1091110305
614 53. Zhang Z, Espy-Wilson CY. 2004 A vocal-tract model of American English/l. *J. Acoust.*
615 *Soc. Am.* 115(3), 1274-1280. https://doi.org/10.1121/1.1645248
616 54. Zhou X, Espy-Wilson CY, Boyce S, Tiede M, Holland C, Choe A. 2008 A magnetic
617 resonance imaging-based articulatory and acoustic study of "retroflex" and
618 "bunched" American English/r. *J. Acous. Soc. Am.* 123(6), 4466-4481.
619 https://doi.org/10.1121%2F1.2902168
620 55. Gamba M, Giacoma C. 2005 Vocal production mechanisms in ruffed lemurs: a
621 prosimian model for the basis of primate phonation. In: *MAVEBA 2005, Proceedings*
622 *of the 4th International Workshop on Models and Analysis of Vocal Emissions for*
623 *Biomedical Applications*; 207-210. Firenze university press.
624 https://doi.org/10.1400/41050
625 56. Gamba M, Medard J, Andriamialison H, Rakotoarisoa G, Giacoma C. 2007 Vocal tract
626 modeling as a tool to investigate species-specific cues in vocalization. In: *Proceedings*
627 *of the 5th International Workshop on Models and Analysis of Vocal Emissions for*
628 *Biomedical Applications*; 225-227. Firenze university press.
629 57. Bates D, Mächler M, Bolker B, Walker S. 2014 Fitting linear mixed-effects models
630 using lme4. arXiv:1406.5823 [stat]. https://doi.org/10.48550/arXiv.1406.5823
631 58. R Core Team 2020. R: a language and environment for statistical computing.
632 Vienna (Austria): R Foundation for Statistical Computing https://www.r-project.org/.
633 59. Estienne V, Mundry R, Kühl HS, Boesch C. 2017 Exploitation of underground bee
634 nests by three sympatric consumers in Loango National Park, Gabon. *Biotropica*. 49,
635 101–109. https://doi.org/10.1111/btp.12354
636 60. Fox J, Monette G. 1992 Generalized collinearity diagnostics. *JASA.* 87(417),
637 178e183. https://doi.org/10.1080/01621459.1992.10475190

638 61. Fox J, Weisberg S. 2011 An R Companion to Applied Regression (2nd ed.) Thousand
639 Oaks, CA: SAGE.
640 62. Akinwande MO, Dikko HG, Samson A. 2015 Variance inflation factor: as a condition
641 for the inclusion of suppressor variable(s) in regression analysis. *Open. J.*
642 *Stat.* 5(07), 754. http://dx.doi.org/10.4236/ojs.2015.57075
643 63. Forstmeier W, Schielzeth H. 2011 Cryptic multiple hypotheses testing in linear
644 models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol.*
645 65, 47–55. https://doi.org/10.1007/s00265-010-1038-5
646 64. Dobson AJ. 2002 An Introduction to Generalized Linear Models. (2nd edn). Boca
647 Raton (FL): Chapman and Hall/CRC Press.
648 65. Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure for
649 confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* 68, 255e278.
650 https://doi.org/10.1016/j.jml.2012.11.001
651 66. Lenth RV. 2016 Least-Squares Means: The R Package lsmeans. *J. Stat.*
652 *Softw. 69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01
653 67. Kazemi A, Kesba M, Provini P. 2023 Realistic three-dimensional avian vocal tract
654 model demonstrates how shape affects sound filtering (*Passer domesticus*). *J. R. Soc.*
655 *Interface.* 20(198), 20220728. https://doi.org/10.1098/rsif.2022.0728
656 68. Riede T, Suthers RA, Fletcher NH, Blevins WE. 2006 Songbirds tune their vocal tract
657 to the fundamental frequency of their song. *PNAS*. 103, 5543–5548.
658 https://doi.org/10.1073/pnas.0601262103
659 69. Riede T, Schilling N, Goller F. 2012 The acoustic effect of vocal tract adjustments in
660 zebra finches. *J. Comp. Physiol. A.* 199, 57–69. https://doi.org/10.1007/s00359-
661 012-0768-4
662 70. Smith DR, Patterson RD, Turner R, Kawahara H, Irino T. 2005 The processing and
663 perception of size information in speech sounds. *J. Acoust. Soc. Am. 117*(1), 305-
664 318. https://doi.org/10.1121/1.1828637
665 71. Puts DA, Hodges CR, Cárdenas RA, Gaulin SJ. 2007 Men's voices as dominance
666 signals: vocal fundamental and formant frequencies influence dominance attributions
667 among men. *Evol. Hum. Behav. 28*(5), 340-344.
668 https://doi.org/10.1016/j.evolhumbehav.2007.05.002
669 72. Charlton B, Reby D, McComb K. 2007 Female red deer prefer the roars of larger
670 males. *Biol. Lett.* 3, 382–385. https://doi.org/10.1098%2Frsbl.2007.0244
671 73. Charlton BD, Zhihe Z, Snyder RJ. 2010. Giant pandas perceive and attend to
672 formant frequency variation in male bleats. Anim. Behav. *79*(6), 1221-1227.

https://doi.org/10.1016/j.anbehav.2010.02.018
74. Fitch WT, Kelley JP. 2000 Perception of vocal tract resonances by whooping cranes Grus americana. *Ethology*. 106, 559-574. https://doi.org/10.1046/j.1439-0310.2000.00572.x
75. Terranova F, Baciadonna L, Maccarone C, Isaja V, Gamba M, Favaro L. 2023 Penguins perceive variations of source- and filter-related vocal parameters of species-specific vocalisations. *Anim. Cogn*. 1-10.
76. Fitch WT. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.* 102(2), 1213-1222. https://doi.org/10.1121/1.421048
77. Collins SA. 2020 Men's voices and women's choices. *Anim. Behav.* 60, 773-780. https://doi.org/10.1006/anbe.2000.1523
78. Charlton BD, Ellis WA, McKinnon AJ, Cowin GJ, Brumm J, Nilsson K, Fitch WT. 2011 Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: honesty in an exaggerated trait. *J. Exp. Biol.* 214(20), 3414-3422. https://doi.org/10.1242/jeb.061358
79. Gonzáles J. 2004 Formant frequencies and body size of speaker: a weak relationship in adult humans. *J. Phon.* 32, 277–287. https://doi.org/10.1016/S0095-4470(03)00049-4
80. Pisanski K, Fraccaro PJ, Tigue CC, O'Connor JJM, Roder S, Andrews PW, Fink B, DeBruine LM, Jones BC, Feinberg DR. 2014 Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* 95, 89-99. https://doi.org/10.1016/j.anbehav.2014.06.011
81. Kriesell HJ, Le Bohec C, Cerwenka AF, Hertel M, Robin JP, Ruthensteiner B, ..., Düring DN. 2020. Vocal tract anatomy of king penguins: morphological traits of two-voiced sound production. *Front. Zool*. 17, 1-11. https://doi.org/10.1186/s12983-020-0351-8
82. Charlton BD, Reby D. 2016 The evolution of acoustic size exaggeration in terrestrial mammals. *Nat. Comm.* 7(1), 12739. https://doi.org/10.1038/ncomms12739
83. Fitch WT. 2000 The evolution of speech: a comparative review. *Trends. Cogn. Sci.* 4, 258267. https://doi.org/10.1016/s1364-6613(00)01494-7