PHILOSOPHICAL
TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Kleisner K *et al*. 2021 Predicting strength from aggressive vocalizations versus speech in African bushland and urban communities. *Phil. Trans. R. Soc. B* **376**: 20200403.
https://doi.org/10.1098/rstb.2020.0403

Accepted: 23 June 2021

One contribution of 11 to a theme issue 'Voice modulation: from origin and mechanism to social impact (Part I)'.

Subject Areas:
behaviour, evolution

Keywords:
nonverbal vocalization, acoustic communication, Hadza, handgrip strength, aggression

Authors for correspondence:
Karel Kleisner
e-mail: karel.kleisner@natur.cuni.cz
Juan David Leongómez
e-mail: leongomez@unbosque.edu.co
Katarzyna Pisanski
e-mail: katarzyna.pisanski@cnrs.fr

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5631308.

THE ROYAL SOCIETY
PUBLISHING

Predicting strength from aggressive vocalizations versus speech in African bushland and urban communities

Karel Kleisner[1], Juan David Leongómez[2], Katarzyna Pisanski[3,4,5], Vojtěch Fiala[1], Clément Cornec[3], Agata Groyecka-Bernard[5], Marina Butovskaya[6,7], David Reby[3], Piotr Sorokowski[5] and Robert Mbe Akoko[8]

[1]Department of Philosophy and History of Science, Charles University, Prague, 12800, Czech Republic
[2]Human Behaviour Lab (LACH), Faculty of Psychology, Universidad El Bosque, Bogota, DC, 110121, Colombia
[3]Equipe de Neuro-Ethologie Sensorielle, Centre de Recherche en Neurosciences de Lyon, Jean Monnet University of Saint-Etienne, 42100, France
[4]CNRS | Centre National de la Recherche Scientifique, Laboratoire Dynamique du Langage, Université Lyon 2, Lyon, 69363, France
[5]Institute of Psychology, University of Wroclaw, 50–527, Poland
[6]Institute of Ethnology and Anthropology, Russian Academy of Science, 119334, Russia
[7]Russian State University for the Humanities, Moscow, 125047, Russia
[8]Department of Communication and Development Studies, University of Bamenda, PO Box 39, Bambili, Bamenda, Cameroon

KK, 0000-0002-3277-8365; JDL, 0000-0002-0092-6298; KP, 0000-0003-0992-2477; VF, 0000-0002-0148-5092; CC, 0000-0002-4110-328X; AG-B, 0000-0002-1932-4828; DR, 0000-0001-9261-1711; PS, 0000-0001-9225-9965

The human voice carries information about a vocalizer's physical strength that listeners can perceive and that may influence mate choice and intrasexual competition. Yet, reliable acoustic correlates of strength in human speech remain unclear. Compared to speech, aggressive nonverbal vocalizations (roars) may function to maximize perceived strength, suggesting that their acoustic structure has been selected to communicate formidability, similar to the vocal threat displays of other animals. Here, we test this prediction in two non-WEIRD African samples: an urban community of Cameroonians and rural nomadic Hadza hunter–gatherers in the Tanzanian bushlands. Participants produced standardized speech and volitional roars and provided handgrip strength measures. Using acoustic analysis and information-theoretic multi-model inference and averaging techniques, we show that strength can be measured from both speech and roars, and as predicted, strength is more reliably gauged from roars than vowels, words or greetings. The acoustic structure of roars explains 40–70% of the variance in actual strength within adults of either sex. However, strength is predicted by multiple acoustic parameters whose combinations vary by sex, sample and vocal type. Thus, while roars may maximally signal strength, more research is needed to uncover consistent and likely interacting acoustic correlates of strength in the human voice.

This article is part of the theme issue 'Voice modulation: from origin and mechanism to social impact (Part I)'.

1. Introduction

Vocalization is among the most powerful communication and signalling channels in all major vertebrate clades [1,2]. Mammalian males of various taxa can produce mighty roars that have a functional role within intrasexual competition and mating contexts, such as the roars of red deer stags that putatively function to exaggerate size and communicate dominance [3]. Humans are no exception as roar-like vocalizations appear within various social interactions such as competition between rivals, combatants and conflicts between larger groups (e.g. sports and warfare [4–7]), as well as in deceptive mimicry of animal calls used for hunting [8]. Multiple converging lines of evidence indicate that nonverbal acoustic features in speech, particularly voice pitch, and more recently the acoustic structure of nonverbal vocalizations such as roars, screams and grunts can influence listeners' judgements of speaker traits and can predict reproductive outcomes, particularly in mate choice and intrasexual competition (see [9] and [10] for review).

Research on the communicative function of agonistic vocalizations in animals has focused largely on the inverse relationship between vocal frequencies, namely fundamental frequency (f_o, perceived as voice pitch) or formant frequencies (resonances of the vocal tract) and body size [11–14]. Indeed, in many mammals including humans, longer vocal tracts result in lower and more closely spaced formants that provide a reliable index of body size [14–17]. While larger individuals with bigger larynges also produce lower f_o, than do smaller individuals [13,18], f_o is not a reliable predictor of body size in many mammals when age and sex are controlled [19], including in humans [17].

A small number of studies have also investigated the vocal correlates of physical strength in the human voice, with mixed and often null results [20–24]. Surprisingly, while human listeners appear capable of assessing the strength of vocalizers from their speech [20,25,26] and from their nonverbal vocalizations, namely roar-like vocalizations (hereafter, 'roars') [25,26], studies have largely failed to find robust and consistent acoustic indices of strength in the human voice. While there is some limited evidence that f_o may predict strength in young peri-pubertal Bolivian Tsimane males after controlling for body size [22], this result has not been consistently replicated in adult men in various other cultures [20,23,26]. Indeed, a recent meta-analysis of eight published studies showed a significant but very weak inverse relationship between adult men's mean f_o and their upper body strength ($r = -0.07$) [24]. Some investigations further suggest a potential link among strength, fighting ability and sexually dimorphic vocal characteristics, but here too the evidence is equivocal for various measures of fighting success [27], hunting reputation [28] and strength [23].

In addition to focusing on a relatively small set of vocal parameters, the vast majority of past research investigating form and function in the human voice has focused almost exclusively on *speech* signals. Importantly, recent research suggests that human roars (compared to speech) may maximize the perceived strength of the vocalizer. Indeed, the acoustic structure of aggressive vocalizations in humans may have been selected to communicate, and to some degree exaggerate, functional cues to physical formidability [25,26], similar to the vocal threat displays of other animals [18,29]. Human roars serve to intimidate the enemy and to motivate an individual or whole group of individuals, as seen in military contexts—in armies as distinctive as the fourteenth century Japanese fighters [4], World War II Soviet soldiers [5] and Ancient Greek troops [6]—and in the context of modern-day sport competitions [7]. Battle cries, and their derivatives in collective sports, are often combined with war dances and intimidating bodily expressions, thus being part of human behavioural patterns of aggression that can lead to the immediate escalation of physical confrontation preceding an attack, or otherwise function to avert it.

Nonverbal vocalizations are also special in human vocal communication because of their acoustic structure: they often occupy a part of acoustic space that corresponds to perceptual roughness and that is separated from other vocal signals, including speech [26,30]. Indeed, unlike speech, human nonverbal vocalizations are often characterized by a high proportion of nonlinear phenomena (NLP) caused by aperiodic vibration of the vocal folds that can give the voice a rough or harsh quality. Nonlinearities are also found in animal distress cries and vocal threat displays [29,31–33], and a wide range of vertebrate species are sensitive to nonlinearities [2,34]. In humans too, they can increase the perceived aversiveness of vocalizations [33]. It has been hypothesized that the nonlinear acoustic parameters that uniquely characterize nonverbal vocalizations may function to communicate physical traits and states, such as physical strength [25,26]. However, given that research on human vocal communication has focused largely on speech, the functional role of nonlinearities in the human voice remains unclear.

Research on the human voice is also largely culturally restricted. While there is relatively rich evidence on the form and function of the human voice (focusing on speech) in Western, educated, industrialized, rich and democratic ('WEIRD') human populations [35], namely European, North American and Asian cultures, less is known about indigenous populations for whom access to media portrayals is limited and, in the case of nomadic tribes, traditional modes of living more closely resemble those of our ancestors. To our knowledge, research on African populations is limited to Namibian Himba and Nama, Cameroonians of Bantu origin, and Tanzanian Hadza [21,28,36–40], and has, with a few exceptions (see [36,40,41]), again focused almost exclusively on speech. In Hadza, males who produce lower-pitched speech have been reported to have more children and, therefore, higher reproductive success [38]. Hunting reputation, a trait associated with reproductive success [42,43] and strength [44], also correlates with lower voice pitch in men and appears to explain much of the variance in reproductive success [28]. Similarly, women whose voices were manipulated to be higher-pitched were regarded as better gatherers [39]. Though the relationship between reproductive success and voice pitch appears negligible when controlling for hunting reputation, hunting reputation remains a reliable predictor of reproductive success when controlling for voice pitch [28]. Both fundamental frequency and handgrip strength (HGS) are positively associated with reproductive success in indigenous Himba women [37].

HGS measured by a dynamometer has been repeatedly shown to predict upper body strength [45,46] and recently, also shown to predict the outcomes of male–male competition in Hadza men [44]. HGS has been used in a series of studies reporting its positive association with men's facial masculinity [47,48], other male sex-specific characteristics and fitness indicators [49–52], and clinical traits associated with health, ageing and mortality (see, e.g. [53–55]). HGS data from rural African and Western populations have revealed comparable relationships with ageing and mortality across populations and thus may represent a cross-culturally robust measure [56]. Although both HGS and sexually dimorphic vocal characteristics are positively correlated with testosterone levels [21,24,57–59], the evidence that individuals with more masculine voices are physically stronger than other individuals of the same sex with less masculine voices remains equivocal [20–24].

Uncovering the vocal correlates of physical strength has emerged as a key question within the evolutionary voice sciences. Previous research assessing physical strength from *speech* has tested both 'WEIRD' vocalizers (namely students from the UK, America or China) and 'non-WEIRD' vocalizers (e.g. Hadza, Bolivian Tsiname) [20–23]. However, only one previous study has tested for acoustic correlates of strength in human nonverbal vocalizations (roars) in a sample of British drama students, but it did not examine the potential role of nonlinear acoustic phenomena (NLP) [26]. In the present study, we take a new approach to this scientific challenge by directly comparing indices of strength in roars to speech in two African non-WEIRD samples and specifically in cultures in which strength is important for evolutionary fitness. We examine the vocalizations of an urban sample of Cameroonian men and women, mainly of Bantu origin, and a rural sample of nomadic Hadza hunter–gatherers living in the Tanzanian bushlands, where access to Western culture including media portrayals of emotional expression is extremely limited and physical strength is important for survival [44,60]. For each individual, voice recordings of standardized speech sentences and volitional aggressive vocalizations in response to hypothetical agonistic contexts were collected, together with measures of HGS. Testing not only speech, but also nonverbal vocalizations (herein, 'roars'), allows us to test the prediction that roars will maximize signals of strength relative to speech, which is far more constrained by the rules of language.

We employ a state-of-the-art information-theoretic approach based on multi-model inference and averaging [61], which we predict may be more powerful than the traditional linear regression models used in previous studies examining vocal indices of strength in the human voice [20–23,26]. Given the exploratory nature of this study, a large number of potential acoustic predictors of strength and the discrepancies and null results of previous studies [20–23], this technique may provide important advantages because it allows the evaluation of multiple models to make inferences without *a priori* selecting one single model as the best approximation to a phenomenon [61], and thus allows us to evaluate the relative importance of numerous nonverbal acoustic parameters in predicting the actual strength of African men and women.

2. Methods

(a) Participants

We collected voice recordings from 141 men and women of African origin, with an even sex ratio, from two distinct samples: urban-dwelling Cameroonians ($N = 101$, 51 women and 50 men) and bush-living Hadza hunter–gatherers ($N = 40$, 20 women and 20 men). The Cameroonians included in this sample were mainly university students of Bantu origin, residing at the time of sampling in the English-speaking Southwest Region. Data were collected at the University of Buea in the regional capital town of Buea. The Hadza are a small nomadic group of hunter–gatherers made up of approximately 1000 individuals residing in small camps in the bushlands of Tanzania [60]. The Hadza share ancestry with Khoisan populations in southern Africa [62]. The distributions of all measured variables (e.g. strength, age and body size) for both men and women are summarized in electronic supplementary material, figure S1, and additional sample descriptions are given in electronic supplementary material, table S1 [63]. All participants provided informed consent to take part in the research and were monetarily compensated for their participation (Cameroonians: 5000 CFA; Hadza: 5000 TZS).

(b) Voice recording

We recorded three types of voice stimuli from each sample of vocalizers—a sentence-long greeting (herein, greeting speech), a series of mono-syllabic vocal sounds or words (vowels or counting, herein, short speech) and volitional nonverbal vocalizations (herein, roar). Cameroonian voice recordings were collected in the Anglophone region of Cameroon (SouthWest) where English is among the official languages and were taken under standardized conditions at the campus of the University of Buea. Vocalizers were recorded individually in a quiet room. Recordings were made using a SONY PCM D50 recorder equipped with a windscreen and saved as WAV files. Each recording session included a short greeting in English: 'Hello, how are you', counting from ten to one (short speech), and three aggressive nonverbal vocalizations (roars). To elicit volitional vocalizations, participants were instructed to imagine themselves in a combat situation in which they are facing the risk of being attacked by an enemy. They were asked to vocally impress the enemy to prevent confrontation or increase their chances of winning a fight, without words, three times in a continuous sequence.

Hadza voice recordings were taken in the Savannah bushland habitat in the Lake Eyasi region of Tanzania. Hadza participants were recorded in individual sessions at 11 different campsites, located with the aid of a local guide and translator who accompanied the field research team on all expeditions. Recordings were taken in a quiet area at an inaudible distance from the campsite to ensure privacy, using a Tascam DR05 recorder equipped by a windscreen and saved as WAV files. Recording sessions followed a standardized procedure in which each participant was instructed to produce a short greeting in Swahili 'Habari gani' (equivalent to 'how are you'), the five vowel sounds /a i e o u/ (short speech) and an aggressive nonverbal vocalization (roar). To elicit volitional vocalizations, participants were instructed to imagine themselves in a combat situation in which they are facing the risk of being attacked by an enemy, and to produce a single nonverbal vocalization, without words, to express threat toward the enemy. They were shown an image of an aggressive war vocalization to aid their interpretation of the task.

(c) Hand grip strength measurement

Using a hand dynamometer, we measured the HGS of each participant's left and right hands three times each. Participants were asked to hold the dynamometer in a vertical position, while standing, with their arm bent at the elbow such that the forearm takes the position perpendicular to the body axis. Participants were further instructed to press the grip of the device with as much force as possible in three rounds for each hand, alternating the right and left hand in each round. Analysis of video-recorded footage of tool use has previously revealed that the Hadza are strikingly lateralized, using the right hand in 96% of all tool-use tasks [64]. Nevertheless, the HGS measures were highly correlated between hands (Hadza: $r = 0.77$; Cameroon: $r = 0.87$) and across three rounds (Hadza: $r = 0.96$; $r = 0.97$; $r = 0.95$; Cameroonian: $r = 0.93$; $r = 0.93$; $r = 0.95$ correlation between first and second, first and third, and second and third HGS measurements, respectively), and thus HGS measures were averaged across all three attempts and for both hands for the purpose of further analysis. We additionally measured participants' weight using metric scales, and height using a portable anthropometer (Hadza) or metric tape affixed to a wall (Cameroon). See figure 1*a* for a graphical representation of HGS distribution and electronic supplementary material, table S1 for HGS descriptive statistics [63].

(d) Acoustic editing and analysis

To standardize recordings between Cameroonian and Hadza samples for acoustic analysis, the greeting vocalizations contained

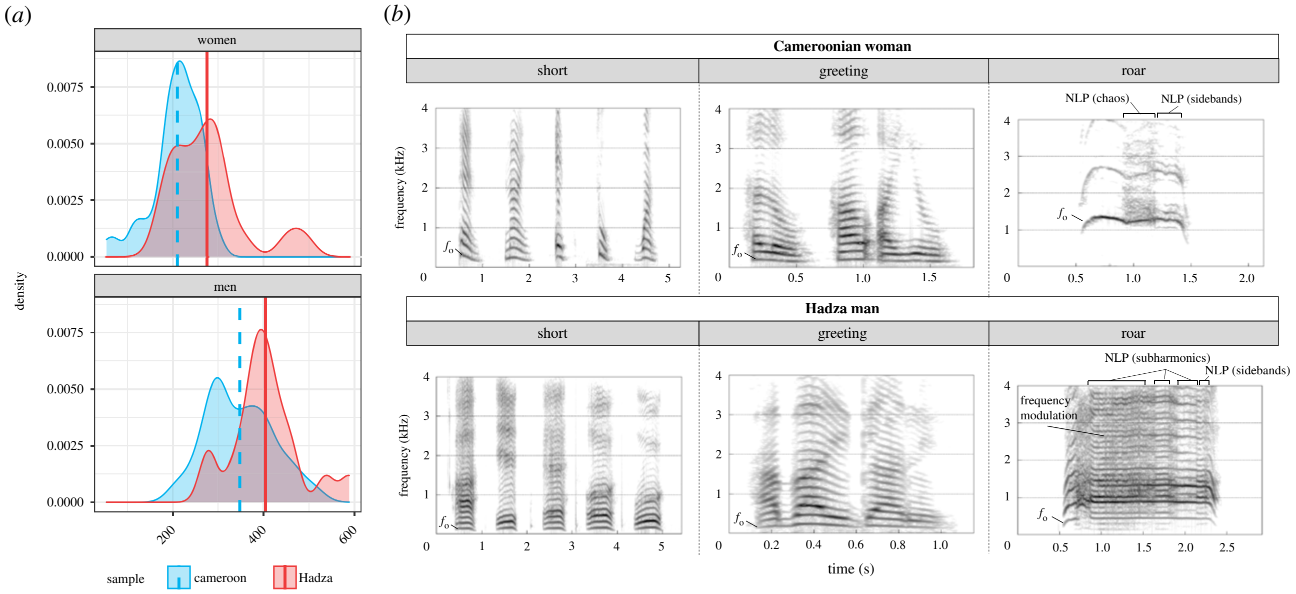

Figure 1. (*a*) Distribution of HGS by sex and sample; coloured vertical lines represent mean HGS. (*b*) Spectrograms of example vocalizations for a Cameroonian woman (top) and a Hadza man (bottom); fundamental frequency (f_o) is labelled in all spectrograms; NLP, including subharmonics and deterministic chaos, as well as frequency modulations (inflex), are common in roars but are largely absent in speech. (Online version in colour.)

five syllables in all cases (i.e. 'He-llo, how-are-you' and 'Ha-ba-ri ga-ni', respectively), and short speech samples were a string of five single-syllable utterances: five vowels in the case of Hadza and five numbers in the case of Cameroonian participants. To account for prosodic changes that mark the beginning and end of a string of utterances, we selected the first three and the last two numbers for Cameroonians (i.e. '10, 9, 8, 2, 1'), and all five vowels for the Hadza (/a i e o u/). For the Cameroonian sample, one of the three roars was selected at random for acoustic analysis. In all cases, recordings were checked for background noise and instances of clipping that could affect acoustic analysis, and we selected recordings with an adequate signal-to-noise ratio. Silence gaps of 0.5 s were inserted at the beginning and end of each recording, and between utterances, for analysis. Voice stimuli thus included '141' greetings, '141' short speech and '135' roars for a total of 417 stimuli used for acoustic analysis. Figure 1*b* provides spectrograms of example vocalizations for each vocal type.

Acoustic analysis was performed in Praat v. 6.1.08 and 6.1.35 [65]. We measured a range of voice parameters (see electronic supplementary material, table S2) predicted to communicate biologically and socially relevant information about a vocalizer, including potentially indicating physical strength ([20,25–27]; see also [9,12] for reviews). These measures included fundamental frequency (f_o) parameters linked to the perception of voice pitch and its variability: f_o mean (figure 1*b*), f_o minimum, and f_o coefficient of variation (f_o CV), mean absolute slope (MAS), and modulation wherein smoothing algorithms were applied to the pitch contour to measure major f_o modulations (inflex2) and minor vibrato-like inflections (inflex25). Fundamental frequency parameters were measured with a search range of 60–2000 Hz, 0.05 s window length and 0.01 time step. Extracted f_o contours were systematically verified and corrected, including where nonlinearities in roar stimuli impeded pitch tracking, following previous work (e.g. [25,66]). We additionally measured vocal perturbation and noise parameters including harmonics-to-noise ratio (HNR, ratio of harmonic to chaotic spectral energy), jitter (minor fluctuations in periodicity) and shimmer (minor fluctuations in amplitude). Finally, we computed the proportion of each roar stimulus containing nonlinear phenomena (%NLP), identified and annotated manually from spectrograms (0–5 kHz; window length 0.05) and amplitude waveforms of voice recordings in Praat (note that only roars contained NLP, our speech sample did not). Proportions of nonlinear vocal phenomena included in %NLP were sidebands (amplitude modulation), subharmonics (vocal fold vibration at an integer fraction of f_o) and deterministic chaos (aperiodic, irregular vocal fold vibration; see [31,33]; see figure 1*b* for examples). With the exception of manually measured nonlinearities, all-acoustic measures were performed using an established custom Praat script (see e.g. [25,26,66]).

See electronic supplementary material, figure S2 for a graphical summary of all acoustic variables and their distributions. Descriptive statistics of the acoustical parameters of Cameroonian and Hadza participants are summarized in the electronic supplementary material, tables S2 and S3, for female and male participants, respectively [63].

(e) Statistical analysis

To analyse the relationship between strength and nonverbal vocal parameters, we employed an information-theoretic approach [61,67,68]. Because this is an exploratory study, by using this technique, we were able to evaluate all possible candidate models derived as subsets of predictors from a global model. In addition, because several candidate models could be equally robust, with no strong reason to prefer one over the others (e.g. candidate models with second-order Akaike Information Criterion ($\Delta AICc$) < 2), we used model averaging techniques to produce both parameter and error estimates derived from weighted averages, and thus not restricted to one model [61].

We built linear models using the *lm* function from base R version 4.1.1 [69]. We fitted separate models for each vocalization type (greeting, short, roar), and for both sexes for a total of six models. We fitted equivalent global models in each case, including sample, mean f_o, minimum f_o, f_o CV, inflex25, inflex2, MAS, HNR, jitter, shimmer and NLP proportion (NLP only for models from roar vocalizations where NLP were present). Individuals with missing data on any model variables were excluded to maintain equal *n*'s, regardless of predictor variables. This step was necessary to ensure that models from different vocalization types were comparable (by *AICc*) within each sample/sex combination. Thus, all models had an equivalent structure and sample size, and were in all cases (with each sample/sex combination) based on vocalizations from the same

Table 1. Information criteria for the best-supported models for each vocal type by sex and sample. *Note:* see electronic supplementary material for detailed information.

	Cameroon					Hadza			
	AICc	*ΔAICc*	*d.f.*	$w_i(AICc)$		*AICc*	*ΔAICc*	*d.f.*	$w_i(AICc)$
women									
roar	465.51	0.00	7	0.95	*greeting*	230.25	0.00	7	0.90
greeting	472.15	6.64	6	0.03	*roar*	235.48	5.23	8	0.07
short	474.31	8.80	3	0.01	*short*	237.15	6.90	4	0.03
men									
roar	486.21	0.00	7	0.91	*short*	209.04	0.00	4	0.50
short	491.70	5.49	5	0.06	*roar*	209.99	0.95	7	0.31
greeting	492.97	6.76	3	0.03	*greeting*	211.04	2.00	3	0.19

participants, with the same data as the dependent predicted variable (strength, HGS).

These highly parametrized global models were subsequently reduced by multi-model inference techniques, using the *dredge* and *model.avg* functions from the R package MuMIn (Multi-Model Inference) [70]. First, the *dredge* function was used to fit all possible combinations (subsets) of predictor terms (both main effects and interactions) from the global model and to rank the generated models based on their *AICc* and Akaike weights ($w_i(AICc)$) [71]. Based on these criteria, we empirically selected the best-supported models (i.e. those with $\Delta AICc \leq 2$ units from the best-supported model) and averaged them using the *model.avg* function. Coefficients were weighted in all averaged models.

To determine which vocalization type best predicts HGS, models for each type of vocalization by sex were then compared in two ways. First, by comparing the top best-supported models using *AICc*, *ΔAICc* and Akaike weights $w_i(AICc)$, and second, by plotting and fitting a linear regression between actual (measured) HGS and the HGS predicted by each averaged model on the basis of vocal parameters. See electronic supplementary material for detailed information on model parameterization, selection, averaging and the related R code.

Given the difference in HGS between the samples (Hadza were significantly stronger than Cameroonian participants; see figure 1*a* and electronic supplementary material, table S1), the sample emerged as a strong predictor in most models, particularly for greeting speech (see electronic supplementary material, figure S11). Thus, to prevent resultant reductions in the predictive power of acoustic variables, we repeated exactly the same modelling, selection and averaging process separately for each combination of sample and sex (thus excluding sample as a predictor in all models). For models of Hadza participants, given the smaller sample size, we limited the number of predictor terms in the subset models produced by *dredge* to between one and 10 (excluding intercept), as recommended by Austin and Steyerberg [72].

3. Results

Multi-model inference procedures first reduced the products of linear modelling, and the best-supported models were subsequently selected. Table 1 presents the information criteria for the top best-supported models for each grouping. The comparison of the best-supported models for predicting strength from each type of vocal stimulus (i.e. single best model per vocal type) revealed a relatively stable pattern for Cameroonians, while the results appeared less consistent among the Hadza (table 1).

Model averaging conveyed clearer results. Figure 2 shows scatterplots resulting from the regression of the actual (measured) strength (HGS) and the strength predicted by each average model, providing a comparison of the predictive power of each voice stimulus type. Roar vocalizations showed the highest predictive power in all examined groups, with the exception of Hadza women. Indeed, the acoustic parameters measured from roars explained the most variance in actual strength for Cameroonian women (40% variance explained), Cameroonian men (41% variance explained) and Hadza men (63% variance explained) (figure 2). In the Cameroonian sample, roars explained more than twice the variance in actual strength than that explained by a short speech, and three to four times more variance than that explained by greeting speech in Cameroonian and Hadza men (figure 2*c*,*d*). Among Hadza women, while roars still explained a high amount of variance in actual strength (71%), models based on roars and greetings had similar predictive power (the difference between coefficients of determination was negligible, $\Delta R^2 = 0.05$, figure 3*b*). In general, the results for Cameroonian participants (both men and women) corroborated the patterns found by the comparisons of the top best-predicted models (table 1), while for Hadza participants, we found only partial congruence between comparisons based on top best-predicted and weighted average models. While this may be due to differences in the vocal expression of strength between Cameroonian and Hadza participants, we cannot rule out the possibility that sample-level differences are due to the smaller sample size for Hadza and thus model instability.

Figure 3 provides coefficient estimates from final average models (models where $\Delta AICc < 2$ were averaged), and thus a detailed summary of the relative consistency and therefore the importance of various acoustic parameters in together predicting the physical strength of vocalizers for each sample of African men and women and for all voice stimulus types. Models for roar vocalizations consistently involved a higher number of acoustic parameters than did models for speech (short or greeting). The most consistently observed predictors of strength across sexes, samples and voice types

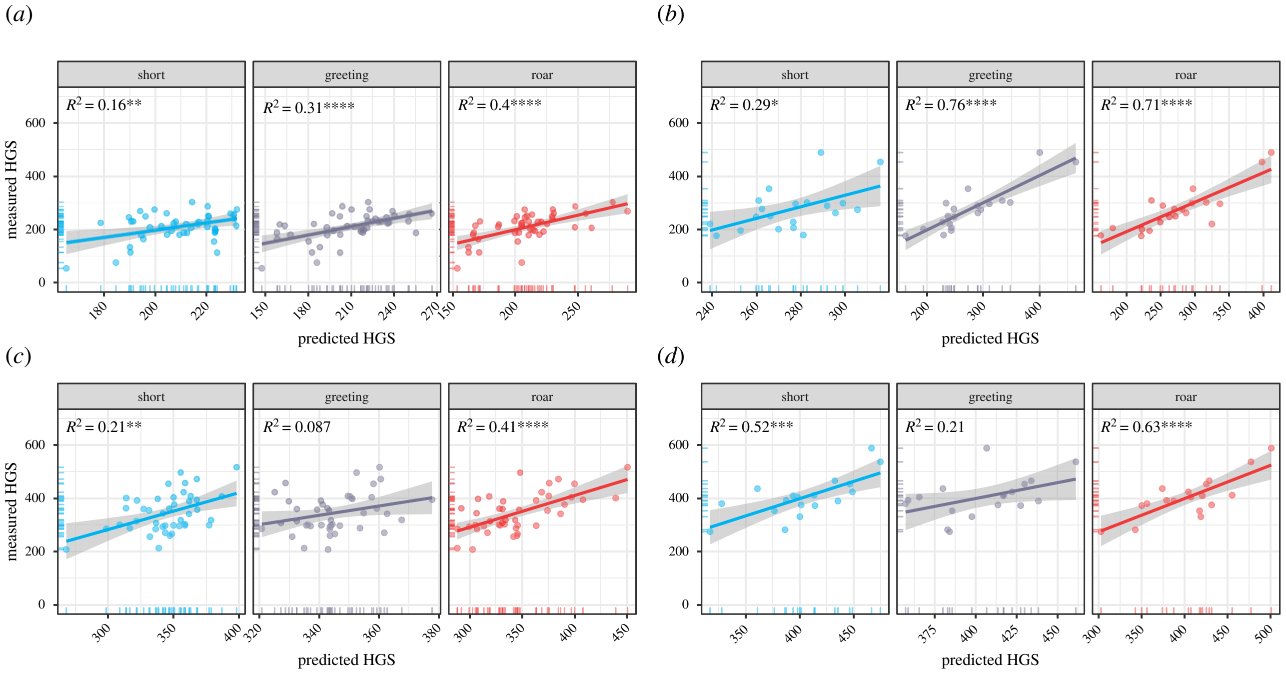

Figure 2. Comparison of the predictive power of the average vocalization models, by vocalization type (*short*, *greeting* and *roar*). (*a*) Cameroonian women, (*b*) Hadza women, (*c*) Cameroonian men and (*d*) Hadza men. Regression lines represent the association between actual hand grip strength (measured HGS, averaged between right and left hand) and the predicted HGS from the final average model from each type of vocalization. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. (Online version in colour.)

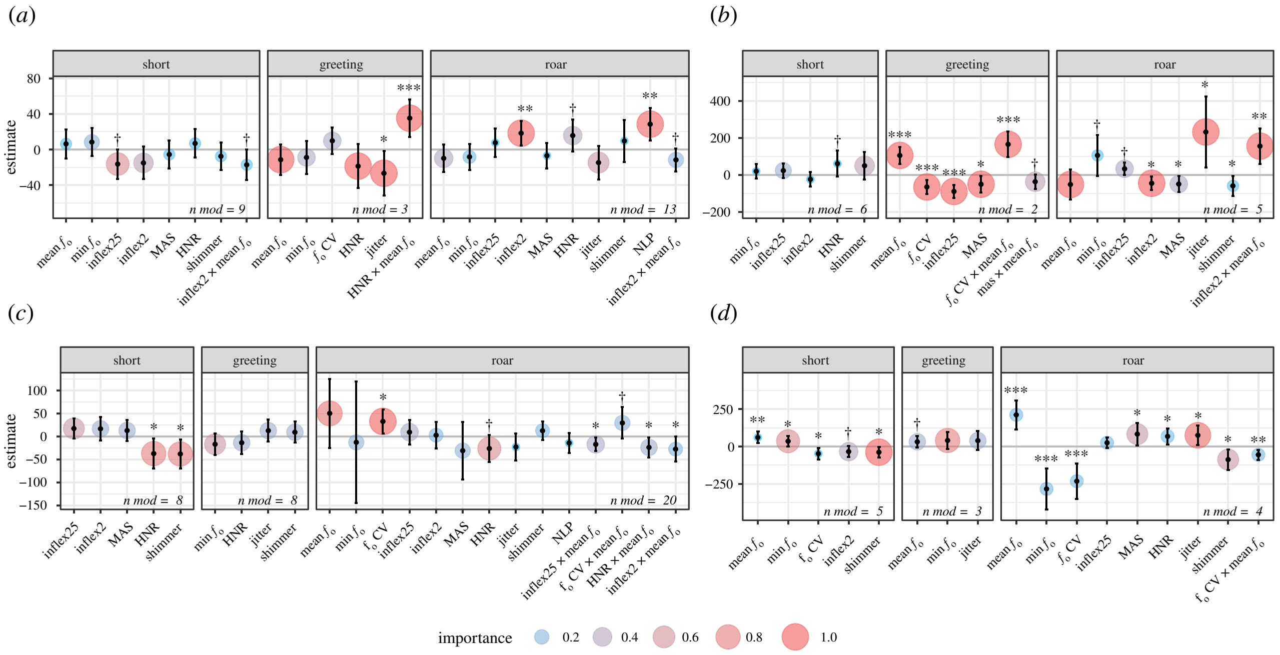

Figure 3. Coefficient estimates of the averaged best models (where *ΔAICc* < 2) by vocalization type (*short*, *greeting* and *roar*). (*a*) Cameroonian women, (*b*) Hadza women, (*c*) Cameroonian men and (*d*) Hadza men. Points and error bars represent average unstandardized weighted *B* estimates and 95% confidence intervals. The colour and size of the bubbles (Importance) represent the proportion of the averaged models in which a term is included, and hence how robust each term is across averaged models. *n mod* is the number of averaged models for each group. HNR = harmonic-to-noise ratio. MAS = mean absolute slope. NLP = nonlinear phenomena. All predictors were centred and scaled. For term significance, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, †$p < 0.10$. (Online version in colour.)

(speech and roars) were mean f_o (voice pitch) and various indices of its variability or instability (f_o CV, jitter, inflex2 and inflex25). However, the direction of these relationships was inconsistent. Indeed, the model averaging results clearly show that vocal indices of strength depend on the combination of multiple acoustic parameters and cannot be predicted as a function of a single acoustic variable, and moreover, that these combinations vary to some extent across voice stimulus type, sex and sample. For example, while nonlinear acoustic phenomena (NLP) appear to play a critical role in predicting strength from the roars of Cameroonian women, along with major modulations in voice pitch (inflex2; figure 3*a*), the contribution of nonlinearities was negligible in the roars produced by men.

Thus, while our results show that human vocal parameters holistically encode information about physical strength and that roars predict strength more effectively than does speech (figure 2), our acoustic modelling techniques (figure 3), like the acoustic analyses of most previous studies ([20,23,26], see also [22]) have failed to identify consistent individual markers of physical strength in the human voice.

4. Discussion

Our results support the prediction that vocal signals to physical strength in humans are maximized in aggressive nonverbal vocalizations (roars) compared to speech. While this prediction has been supported in a Western population (UK drama students: [25,26]), here we extend this research to two African samples, one from the relatively urbanized municipality of Buea (students at the local university) and the other from a rural and nomadic small-scale population of Hadza hunter–gatherers. Applying a bottom-up information-theoretic modelling approach, we show that the nonverbal acoustic structure of roars best predicts physical strength. Indeed, predicted strength based on vocal parameters in roars explained the most variance in actual strength for Cameroonian men and women (explaining 40% of the variance in measured HGS) and for Hadza men (explaining 63% of the variance), and explained generally two to four times more variance in strength than did speech (vowels, words or phrases). While roars relative to greetings predicted strength better in men than in women, roars produced by Hadza women explained an impressive 71% of the variance in their actual physical strength, though this was comparable to the predictive power of their greeting speech (77%). Thus, in contrast with speech, nonverbal roars appear to most effectively encode functional cues to physical strength, as also observed in non-human mammals [29].

However, despite our finding that roars and, to a lesser extent, speech, encode information about physical strength in non-WEIRD samples of men and women of African origin, our analyses did not identify a single vocal parameter nor a consistent combination of vocal parameters that predicted strength in both sexes and in both speech and roars. The complex combinations of acoustic predictors revealed by our models, and their high variability across sex, sample and vocal stimulus type, corroborate the discrepancies of past studies conducted in Western samples [20,22–24,26].

In an attempt to overcome the mixed and null results of this past work, we (i) employed an information-theoretic approach [61,67,68] in order to more extensively explore potential acoustic predictors of strength; (ii) examined these predictors in both speech and roars, wherein the latter was predicted to carry more information about physical formidability [25,26]; and (iii) tested for acoustic indices of strength in two non-WEIRD African samples. In both samples, but particularly among the Hadza, physical strength may significantly contribute to the biological fitness of an individual given that it positively affects hunting outcomes [44]. Therefore, acoustic communication may be an optimal way to mediate social dominance hierarchies and maintain resource-control without engaging in a risky physical confrontation. Indeed, we found that Hadza men and women were physically stronger than our more urban sample of Cameroonian men and women (on average by 16–31%) and that roars predicted strength better in Hadza men and women than in Cameroonian men and women. However, we also found that acoustic predictors of actual strength were more difficult to identify and less stable in the Hadza sample. The reasons for this could be ecological. For instance, Hadza are bush-living people who often communicate at long distances using loud vocalizations or speech, whereas our Cameroonian sample is urbanized, and more often communicate at shorter distances and at a lower volume. The two samples also speak different languages. While Cameroonians from Southwest and Northwest regions speak fluent English, alongside a variety of local native languages, the Hadza speak Swahili and/or Hadzane, a click language consisting of three types of click consonants that may be produced in voiceless oral, voiced nasal, or voiceless nasal, and glottalised variant [60]. Despite these differences, we cannot rule out the possibility that sample-level differences emerged due to small sample size in the Hadza. Indeed the small sample size of the Hadza is a key limitation of this study. While data from extreme non-WEIRD samples are rare and difficult to obtain, the small sample size may have contributed to inconsistencies in the predictive power of vocal parameters and these results thus should be interpreted with caution.

Regarding specific acoustic parameters, it is difficult to derive a clear generalization of their independent contributions due to the lack of consistency in the pattern of acoustic predictors included in each final average model. However, unlike in studies based on assessments of formidability in voice perception (e.g. [73]), and evidence that relatively low f_o can predict strength in the speech of peri-pubertal Bolivian Tsiname males (but not females; [22]), we did not find a consistent relationship between low male fundamental frequency (f_o) and strength across samples and different vocal types. In fact, in several cases, for example in the short speech and roars of Hadza men, *higher* mean f_o signalled strength. As increased subglottal pressure will cause an increase in voice pitch [74], this result could be due to greater lung capacity and/or louder vocalizations produced by stronger men, a prediction that can be directly tested in future work. Notably, a recent meta-analysis showed, using data from eight studies and 845 adult men, that mean f_o explains a mere 0.005% of the variance ($r = -0.07$) in men's upper body strength [24]. The present study is, to our knowledge, the first to examine whether nonlinear acoustic phenomena (NLP) predict strength in human roars. While we find preliminary evidence to support this, the positive relationship between NLP and strength was most evident in Cameroonian women's roars. In order to reduce the number of terms in our statistical models, we computed a single cumulative proportion (%NLP) combining sidebands, subharmonics and deterministic chaos. This cumulative proportion has previously been shown to reliably index ostensible pain levels in volitional human pain vocalizations [75]. However, we cannot rule out the possibility that specific NLP sub-types (e.g. deterministic chaos, which is typically the most strongly associated with affective intensity [33]) may predict strength more effectively than others. This possibility can be tested in future studies that employ larger samples of vocalizers to ensure adequate sampling of various sub-types of NLP in nonverbal vocalizations, and adequate statistical power to test their relative roles.

5. Conclusion

Altogether our findings offer four key conclusions. First, we replicate the finding that both speech and roars can predict strength, and critically, that roars tend to predict strength better than does speech. Moreover, we extend previous work with UK students [25,26] to show that this is true across two additional extremely different human cultures (Cameroonian urban-dwellers and Hadza hunter–gatherers). Second, we show that roars consistently predict HGS with a relatively higher level of accuracy than does speech, explaining 40–80% of the variance in strength within adults of either sex (association between actual and predicted HGS, $R^2 > 0.4$ in all cases; see figure 2). Third, our acoustic analyses show that strength cannot be predicted solely by one acoustic variable, such as voice pitch or NLP alone; rather it seems that vocal indices of strength are likely to depend on combinations of multiple acoustic parameters. Thus, to uncover the clearly complex vocal predictors of physical strength, our results suggest that researchers may need to employ likewise complex models with multiple predictors and interaction terms. In addition, having larger sample sizes and a broader range of samples from different cultural and linguistic backgrounds (at least five) will allow researchers to include those samples as levels in a random factor, with random intercepts as well as random slopes for specific acoustic characteristics. This may provide a clearer picture of the effect of individual voice parameters in communicating strength across diverse human cultures and could produce more generalizable results. Fourth, our results show that volitionally produced human roars retain honest information about a vocalizer's actual physical strength. This research thus adds to a growing body of literature examining form and function in human vocalizations in light of the special capacity of humans to volitionally modulate our vocal output or to produce vocalizations entirely on demand [76–78]. Such a capacity for volitional voice modulation in humans is not observed to the same extent in other mammals, including other primates [76]. In addition to being a precursor to articulated speech [78,79], the capacity to modulate our voices could be beneficial for our fitness, for instance in the context of deceptive signalling of body size [80], and strength [25,26], an important avenue for continued research.

Ethics. Experiment protocols involving humans were in accordance with the guidelines of the Declaration of Helsinki and followed all national/international/institutional guidelines. Procedures were approved by the Institutional Review Board of the Faculty of Science of the Charles University (protocol ref. no. 06/2017) and by the Scientific Council of the Institute of Ethnology, Russian Academy of Sciences, protocol N 1., issued on 19th February 2015. Research on the Hadza population was approved by the Tanzania Commission for Science and Technology (2016-176-ER-2009-151).

Data accessibility. Electronic supplementary material, including code and data, are available online at: https://doi.org/10.17605/OSF.IO/JU6M8.

Authors' contributions. K.K., K.P., J.D.L., D.R. and V.F. conceived and designed the research studies; K.K., K.P., J.D.L. and V.R. conceived the analytical approach and drafted the manuscript; J.D.L., K.P. and C.C. performed acoustic analyses; J.D.L. performed statistical analyses, wrote the code and created supplementary materials; J.D.L. and C.C. designed the figures; K.K. and R.M.A. collected field data on the Cameroonian population; K.P., A.G.B., M.B. and P.S. collected field data on the Hadza population; all authors critically reviewed and commented on the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Competing interests. The authors declare no competing interests.

Funding. Part of this research was supported by a research grant from the National Science Centre to P.S. A.G.-B. and K.P. (grant no. 2016/23/B/HS6/00771). J.D.L was supported by Universidad El Bosque, Vice-rectory of Research (grant no. PCI.2015-8207). K.K. and V.F. were supported by the Czech Science Foundation (grant no. 18-10298S) and by the Charles University Research Centre program no. 204056. D.R., C.C. and K.P were supported by the University of Lyon IDEXLyon project as part of the 'Programme Investissements d'Avenir' (grant no. ANR-16-IDEX-0005). A.G.B was supported by Foundation for Polish Science (scholarship FNP START 2021).

Acknowledgements. M.B. conducted the study in line with the research plans of the Institute of Ethnology and Anthropology RAS. The authors thank all Cameroonians from the University of Buea and Hadza men and women for their participation in this study, and thank the guides and translators for field research assistance.

References

1. Zelick R, Mann DA, Popper AN. 1999 Acoustic communication in fishes and frogs. In *Comparative hearing: fish and amphibians* (eds RR Fay, AN Popper), pp. 363–411. New York, NY: Springer.
2. Ladich F, Winkler H. 2017 Acoustic communication in terrestrial and aquatic vertebrates. *J. Exp. Biol.* **220**, 2306–2317. (doi:10.1242/jeb.132944)
3. 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. (doi:10.1006/anbe.2003.2078)
4. Conlan T. 1999 The nature of warfare in Fourteenth-Century Japan: the record of Nomoto Tomoyuki. *J. Jpn. Stud.* **25**, 299–330. (doi:10.4324/9781315234328-24)
5. Merridale C. 2006 Culture, ideology and combat in the Red Army, 1939–45. *J. Contemp. Hist.* **41**, 305–324. (doi:10.1177/0022009406062072)
6. Burkert W. 1992 *The orientalizing revolution: near eastern influences on Greek culture in the early archaic Age*, 2nd edn. Cambridge, MA: Harvard University Press.
7. Šebesta P, Třebický V, Fialová J, Havlíček J. 2019 Roar of a champion: loudness and voice pitch predict perceived fighting ability in MMA fighters. *Front. Psychol.* **10**, 859. (doi:10.3389/fpsyg.2019.00859)
8. Knight C, Lewis J. 2017 Wild voices: mimicry, reversal, metaphor, and the emergence of language. *Curr. Anthropol.* **58**, 435–453. (doi:10.1086/692905)
9. Pisanski K, Bryant GA. 2019 The evolution of voice perception. In *The Oxford handbook of voice studies* (eds NS Eidsheim, K Meizel), pp. 268–300. New York, NY: Oxford University Press.
10. Puts D. 2016 Human sexual selection. *Curr. Opin. Psychol.* **7**, 28–32. (doi:10.1016/j.copsyc.2015.07.011)
11. Bowling DL, Garcia M, Dunn JC, Ruprecht R, Stewart A, Frommolt KH, Fitch WT. 2017 Body size and vocalization in primates and carnivores. *Sci. Rep.* **7**, 41070. (doi:10.1038/srep41070)
12. Charlton BD, Pisanski K, Raine J, Reby D. 2020 Coding of static information in terrestrial mammal vocal signals. In *Coding strategies in vertebrate acoustic communication* (eds T Aubin, M Mathevon), pp. 115–136. Cham, Switzerland: Springer.
13. Fitch WT, Hauser MD. 2003 Unpacking 'Honesty': vertebrate vocal production and the evolution of acoustic signals. In *Acoustic communication*, (eds AM Simmons, RR Fay, AN Popper) pp. 65–137. New York, NY: Springer. (doi:10.1007/0-387-22762-8_3)
14. Fitch WT. 2000 The phonetic potential of nonhuman vocal tracts: comparative cineradiographic

observations of vocalizing animals. *Phonetica* **57**, 205–218. (doi:10.1159/000028474)

15. Fitch WT. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoust. Soc. Am.* **102**, 1213–1222. (doi:10.1121/1.421048)
16. Fitch WT, Reby D. 2001 The descended larynx is not uniquely human. *Proc. R. Soc. B* **268**, 1669–1675. (doi:10.1098/rspb.2001.1704)
17. Pisanski K *et al*. 2014 Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* **95**, 89–99. (doi:10.1016/j.anbehav.2014.06.011)
18. Morton ES. 1977 On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869. (doi:10.1086/283219)
19. Charlton BD, Reby D. 2016 The evolution of acoustic size exaggeration in terrestrial mammals. *Nat. Commun.* **7**, 12739. (doi:10.1038/ncomms12739)
20. Sell A, Bryant GA, Cosmides L, Tooby J, Sznycer D, Von Rueden C, Krauss A, Gurven M. 2010 Adaptations in humans for assessing physical strength from the voice. *Proc. R. Soc. B* **277**, 3509–3518. (doi:10.1098/rspb.2010.0769)
21. Puts DA, Apicella CL, Cárdenas RA. 2012 Masculine voices signal men's threat potential in forager and industrial societies. *Proc. R. Soc. B* **279**, 601–609. (doi:10.1098/rspb.2011.0829)
22. Hodges-Simeon CR, Gurven M, Puts DA, Gaulin SJCC. 2014 Vocal fundamental and formant frequencies are honest signals of threat potential in peripubertal males. *Behav. Ecol.* **25**, 984–988. (doi:10.1093/beheco/aru081)
23. Han C, Wang H, Fasolt V, Hahn AC, Holzleitner IJ, Lao J, DeBruine LM, Feinberg DR, Jones BC. 2018 No clear evidence for correlations between handgrip strength and sexually dimorphic acoustic properties of voices. *Am. J. Hum. Biol.* **30**, 1–4. (doi:10.1002/ajhb.23178)
24. Aung T, Puts D. 2020 Voice pitch: a window into the communication of social power. *Curr. Opin. Psychol.* **33**, 154–161. (doi:10.1016/j.copsyc.2019.07.028)
25. Raine J, Pisanski K, Oleszkiewicz A, Simner J, Reby D. 2018 Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* **4**, 273–280. (doi:10.1016/j.isci.2018.05.002)
26. Raine J, Pisanski K, Bond R, Simner J, Reby D. 2019 Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech. *PLoS ONE* **14**, e0213034. (doi:10.1371/journal.pone.0213034)
27. Aung T, Goetz S, Adams J, McKenna C, Hess C, Roytman S, Cheng JT, Zilioli S, Puts DA. 2021 Low fundamental and formant frequencies predict fighting ability among male mixed martial arts fighters. *Sci. Rep.* **11**, 905. (doi:10.1038/s41598-020-79408-6)
28. Smith KM, Olkhov YM, Puts DA, Apicella CL. 2017 Hadza men with lower voice pitch have a better hunting reputation. *Evol. Psychol.* **15**, 1–12. (doi:10.1177/1474704917740466)
29. Frey R, Gebler A. 2010 Mechanisms and evolution of roaring-like vocalization in mammals. In *Handbook of behavioral neuroscience* **19** (ed. Stefan M. Brudzynski), pp. 439–450. San Diego, CA: Elsevier B.V. (doi:10.1016/B978-0-12-374593-4.00040-1)
30. Arnal LH, Flinker A, Kleinschmidt A, Giraud AL, Poeppel D. 2015 Human screams occupy a privileged niche in the communication soundscape. *Curr. Biol.* **25**, 2051–2056. (doi:10.1016/j.cub.2015.06.043)
31. Fitch WT, Neubauer J, Herzel H. 2002 Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Anim. Behav.* **63**, 407–418. (doi:10.1006/anbe.2001.1912)
32. Karp D, Manser MB, Wiley EM, Townsend SW. 2014 Nonlinearities in Meerkat alarm calls prevent receivers from habituating. *Ethology* **120**, 189–196. (doi:10.1111/eth.12195)
33. Anikin A, Pisanski K, Reby D. 2020 Do nonlinear vocal phenomena signal negative valence or high emotion intensity? *R. Soc. Open Sci.* **7**, 201306. (doi:10.1098/rsos.201306)
34. Herbst CT. 2016 Biophysics of vocal production in mammals. In *Vertebrate sound production and acoustic communication* (eds RA Suthers, WT Fitch, RR Fay, AN Popper), pp. 159–189. Cham, Switzerland: Springer.
35. Henrich J, Heine SJ, Norenzayan A. 2010 Most people are not WEIRD. *Nature* **466**, 29. (doi:10.1038/466029a)
36. Sauter DA, Eisner F, Ekman P, Scott SK. 2010 Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl Acad. Sci. USA* **107**, 2408–2412. (doi:10.1073/pnas.0908239106)
37. Atkinson J, Pipitone RN, Sorokowska A, Sorokowski P, Mberira M, Bartels A, Gallup GGJ. 2012 Voice and handgrip strength predict reproductive success in a group of indigenous African females. *PLoS ONE* **7**, e41811. (doi:10.1371/journal.pone.0041811)
38. Apicella CL, Feinberg DR, Marlowe FW. 2007 Voice pitch predicts reproductive success in male hunter-gatherers. *Biol. Lett.* **3**, 682–684. (doi:10.1098/rsbl.2007.0410)
39. Apicella CL, Feinberg DR. 2009 Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proc. R. Soc. B* **276**, 1077–1082. (doi:10.1098/rspb.2008.1542)
40. Šebesta P, Kleisner K, Tureček P, Kočnar T, Akoko RM, Třebický V, Havlíček J. 2017 Voices of Africa: acoustic predictors of human male vocal attractiveness. *Anim. Behav.* **127**, 205–211. (doi:10.1016/j.anbehav.2017.03.014)
41. Bryant GA *et al*. 2016 Detecting affiliation in colaughter across 24 societies. *Proc. Natl Acad. Sci. USA* **113**, 4682–4687. (doi:10.1073/pnas.1524993113)
42. Marlowe FW. 2001 Male contribution to diet and female reproductive success among foragers. *Curr. Anthropol.* **42**, 755–759. (doi:10.1086/323820)
43. Apicella CL. 2014 Upper-body strength predicts hunting reputation and reproductive success in Hadza hunter–gatherers. *Evol. Hum. Behav.* **35**, 508–518. (doi:10.1016/j.evolhumbehav.2014.07.001)
44. Misiak M, Butovskaya ML, Oleszkiewicz A, Sorokowski P. 2020 Digit ratio and hand grip strength are associated with male competition outcomes: a study among traditional populations of the Yali and Hadza. *Am. J. Hum. Biol.* **32**, e23321. (doi:10.1002/ajhb.23321)
45. Trosclair D, Bellar D, Judge LWW, Smith J, Mazerat N, Brignac A. 2011 Hand-grip stength as a predictor of muscular strength and endurance. *J. Strength Cond. Res.* **25**, S99. (doi:10.1097/01.JSC.0000395736.42557.bc)
46. Gallup AC, Fink B. 2018 Handgrip strength as a Darwinian fitness indicator in men. *Front. Psychol.* **9**, 439. (doi:10.3389/fpsyg.2018.00439)
47. Fink B, Neave N, Seydel H. 2007 Male facial appearance signals physical strength to women. *Am. J. Hum. Biol.* **19**, 82–87. (doi:10.1002/ajhb.20583)
48. Windhager S, Schaefer K, Fink B. 2011 Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *Am. J. Hum. Biol.* **23**, 805–814. (doi:10.1002/ajhb.21219)
49. Gallup AC, O'Brien DT, White DD, Wilson DS. 2010 Handgrip strength and socially dominant behavior in male adolescents. *Evol. Psychol.* **8**, 229–243. (doi:10.1177/147470491000800207)
50. Gallup AC, White DD, Gallup GGJ. 2007 Handgrip strength predicts sexual behavior, body morphology, and aggression in male college students. *Evol. Hum. Behav.* **28**, 423–429. (doi:10.1016/j.evolhumbehav.2007.07.001)
51. Shoup ML, Gallup GGJ. 2008 Men's faces convey information about their bodies and their behavior: what you see is what you get. *Evol. Psychol.* **6**, 469–479. (doi:10.1177/147470490800600311)
52. Young RW. 2003 Evolution of the human hand: the role of throwing and clubbing. *J. Anat.* **202**, 165–174. (doi:10.1046/j.1469-7580.2003.00144.x)
53. McGrath RP, Johnson N, Klawitter L, Mahoney S, Trautman K, Carlson C, Rockstad E, Hackney KJ. 2020 What are the association patterns between handgrip strength and adverse health conditions? A topical review. *SAGE Open Med.* **8**, 1–12. (doi:10.1177/2050312120910358)
54. Smith L, Yang L, Hamer M. 2019 Handgrip strength, inflammatory markers, and mortality. *Scand. J. Med. Sci. Sport.* **29**, 1190–1196. (doi:10.1111/sms.13433)
55. McGrath RP, Kraemer WJ, Snih SAl, Peterson MD. 2018 Handgrip strength and health in aging adults. *Sport. Med.* **48**, 1993–2000. (doi:10.1007/s40279-018-0952-y)
56. Koopman JJE, van Bodegom D, van Heemst D, Westendorp RGJ. 2015 Handgrip strength, ageing and mortality in rural Africa. *Age Ageing* **44**, 465–470. (doi:10.1093/ageing/afu165)

57. Dabbs JM, Mallinger A. 1999 High testosterone levels predict low voice pitch among men. *Pers. Individ. Dif.* **27**, 801–804. (doi:10.1016/S0191-8869(98)00272-4)
58. Evans S, Neave N, Wakelin D, Hamilton C. 2008 The relationship between testosterone and vocal frequencies in human males. *Physiol. Behav.* **93**, 783–788. (doi:10.1016/j.physbeh.2007.11.033)
59. Butovskaya ML, Windhager S, Karelin D, Mezentseva A, Schaefer K, Fink B. 2018 Associations of physical strength with facial shape in an African pastoralist society, the Maasai of Northern Tanzania. *PLoS ONE* **13**, e0197738. (doi:10.1371/journal.pone.0197738)
60. Marlowe FW. 2010 *The Hadza: hunter–gatherers of Tanzania*, 3rd edn. Berkeley, CA: University of California Press. See https://www.jstor.org/stable/10.1525/j.ctt1pp17z.
61. Symonds MREE, Moussalli A. 2011 A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behav. Ecol. Sociobiol.* **65**, 13–21. (doi:10.1007/s00265-010-1037-6)
62. Shriner D, Tekola-Ayele F, Adeyemo A, Rotimi CN. 2018 Genetic ancestry of Hadza and Sandawe peoples reveals ancient population structure in Africa. *Genome Biol. Evol.* **10**, 875–882. (doi:10.1093/gbe/evy051)
63. Kleisner K *et al.* 2021 Data and code from: Predicting strength from aggressive vocalisations versus speech in African bushland and urban communities. [Database]. *Open Science Framework*. (doi:10.17605/OSF.IO/JU6M8)
64. Cavanagh T, Berbesque JC, Wood B, Marlowe FW. 2016 Hadza handedness: lateralized behaviors in a contemporary hunter–gatherer population. *Evol. Hum. Behav.* **37**, 202–209. (doi:10.1016/j.evolhumbehav.2015.11.002)
65. Boersma P, Weenink D. 2020 Praat: doing phonetics by computer. [computer program]. (See https://www.fon.hum.uva.nl/praat/)
66. Pisanski K, Raine J, Reby D. 2020 Individual differences in human voice pitch are preserved from speech to screams, roars and pain cries. *R. Soc. Open Sci.* **7**, 191642. (doi:10.1098/rsos.191642)
67. Burnham KP, Anderson DR, Huyvaert KP. 2011 AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav. Ecol. Sociobiol.* **65**, 23–35. (doi:10.1007/s00265-010-1029-6)
68. Johnson JB, Omland KS. 2004 Model selection in ecology and evolution. *Trends Ecol. Evol.* **19**, 101–108. (doi:10.1016/j.tree.2003.10.013)
69. R Core Team. 2021 R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. (https://www.R-project.org/)
70. Bartón K. 2020 MuMIn: multi-model inference. R package version 1.43.17. (See https://rdrr.io/cran/MuMIn/.)
71. Wagenmakers EJ, Farrell S. 2004 AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196. (doi:10.3758/bf03206482)
72. Austin PC, Steyerberg EW. 2015 The number of subjects per variable required in linear regression analyses. *J. Clin. Epidemiol.* **68**, 627–636. (doi:10.1016/j.jclinepi.2014.12.014)
73. Hodges-Simeon CR, Gaulin SJCC, Puts DA. 2010 Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* **21**, 406–427. (doi:10.1007/s12110-010-9101-5)
74. Titze IR. 1994 *Principles of voice production*. 1st edn. Englewood Cliff, NJ: Prentice Hall.
75. Raine J, Pisanski K, Simner J, Reby D. 2019 Vocal communication of simulated pain. *Bioacoustics* **28**, 404–426. (doi:10.1080/09524622.2018.1463295)
76. Ackermann H, Hage SR, Ziegler W. 2014 Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* **37**, 529–604. (doi:10.1017/S0140525X13003099)
77. Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. 2016 Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* **20**, 304–318. (doi:10.1016/j.tics.2016.01.002)
78. Fitch WT. 2018 The biology and evolution of speech: a comparative analysis. *Annu. Rev. Linguist.* **4**, 255–279. (doi:10.1146/annurev-linguistics-011817-045748)
79. Boë LJ, Sawallis TR, Fagot J, Badin P, Barbier G, Captier G, Ménard L, Heim JL, Schwartz JL. 2019 Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Sci. Adv.* **5**, eaaw3916. (doi:10.1126/sciadv.aaw3916)
80. Pisanski K, Reby D. 2021 Efficacy in deceptive vocal exaggeration of human body size. *Nat. Commun.* **12**, 968. (doi:10.1038/s41467-021-21008-7)