

Efficacy in deceptive vocal exaggeration of human body size

Katarzyna Pisanski ^{1,2}✉ & David Reby ¹

How can deceptive communication signals exist in an evolutionarily stable signalling system? To resolve this age-old honest signalling paradox, researchers must first establish whether deception benefits deceivers. However, while vocal exaggeration is widespread in the animal kingdom and assumably adaptive, its effectiveness in biasing listeners has not been established. Here, we show that human listeners can detect deceptive vocal signals produced by vocalisers who volitionally shift their voice frequencies to exaggerate or attenuate their perceived size. Listeners can also judge the relative heights of cheaters, whose deceptive signals retain reliable acoustic cues to interindividual height. Importantly, although vocal deception biases listeners' absolute height judgments, listeners recalibrate their height assessments for vocalisers they correctly and concurrently identify as deceptive, particularly men judging men. Thus, while size exaggeration can fool listeners, benefiting the deceiver, its detection can reduce bias and mitigate costs for listeners, underscoring an unremitting arms-race between signallers and receivers in animal communication.

¹Equipe de Neuro-Ethologie Sensorielle (ENES), Centre de Recherche en Neurosciences de Lyon (CRNL), CNRS, INSERM, University of Lyon/Saint-Étienne, Saint-Étienne, France. ²Institute of Psychology, University of Wrocław, Wrocław, Poland. ✉email: kasiapisanski@gmail.com

The honest signalling paradox has perplexed researchers in animal communication for five decades^{1–9}. Indeed, why should receivers pay attention to deceptive signals if these signals contain little to no reliable information? Conversely, if receivers detect and correct for deception, why should signallers continue to produce potentially costly deceptive displays?

At the crux of this paradox lies the inherent conflict of interest between signallers and receivers. While earlier theories saw animal communication as a cooperative exchange of information¹⁰, the production of animal signals during communication between unrelated individuals is now predominantly regarded as a selfish behaviour^{2,10}, whereby signallers attempt to manipulate the responses of receivers. Crucially, however, selection also operates on receivers to evade deception, for example by leading them to ignore deceptive signals or to recalibrate their responses. Indeed both individuals in a dyadic exchange are expected to behave in ways that maximise their own fitness², giving rise to an evolutionary arms race that is most apparent when interests diverge (e.g., mate choice) or are entirely opposed (e.g., resource contests)¹¹. Yet, even in ostensibly cooperative contexts such as alarm calling¹², signallers may stand to gain substantial fitness benefits by exaggerating (or in some cases, attenuating¹³) signals, if such deceptive signals succeed to elicit a beneficial response from receivers^{6,7}.

To remain ‘evolutionarily stable’¹⁴, signals must confer net fitness benefits to both senders and receivers², and thus should be generally reliable or ‘honest on average’^{4,15}. Even a putatively deceptive signal should be reliable enough to remain beneficial for receivers to attend to it. In other words, not only can deception and reliability coexist, deception depends on reliability, because without some element of truth a signalling system would collapse⁶. Signal reliability can be imposed by a number of mechanisms, including anatomical or physiological constraints (e.g., by-product information¹⁶ or honest indices¹⁵), developmental or metabolic costs^{1,17}, and reputation or retaliation costs^{9,18}. Constraints and costs, if high enough, can enforce signal honesty. Nevertheless, even a presumably reliable signal will be ‘incompletely honest’⁷ due either to deceptive processes that must operate within these constraints, including anatomical

adaptations to the vocal apparatus in many mammals¹⁹, or nondeceptive processes, including developmental noise²⁰, communication errors, or environmental signal degradation⁷. While challenging⁶, it is imperative to dissociate deceptive from non-deceptive processes that can independently degrade signal reliability if we are to truly understand the evolution of a signalling system⁷.

Signal reliability is often measured as the strength of the correlation between the signal and the intended information^{6,7}. To gauge signal reliability we must establish (a) if the signal is reliable enough that a receiver will generally benefit by attending to it; (b) the constraints or costs that impose this degree of honesty; and perhaps most critically, (c) how receivers respond to the signal, including detecting and compensating, at least in part, for any deception⁶. Two major obstacles have limited the extent to which we can answer these questions using animal models: uncertainty in what a signal is actually intended to convey and uncertainty in what a receiver is attending to⁶.

We propose that studying deception in human communication signals offers a promising solution to these long-standing hurdles. Unlike other animals, humans can produce specific deceptive signals on demand^{9,21–23}, thus eliminating uncertainty in the signal’s intended function and allowing researchers to pin-point the contribution of deception in signal reliability. Moreover, researchers can directly measure the effects of this deception on human receivers using controlled psychoacoustic experiments, offering a full picture of the signaller–receiver communication chain (Fig. 1). Here, we apply this paradigm to study an ecologically relevant vocal signalling system observed across the animal kingdom—vocal communication of body size.

Given the often substantial fitness benefits of a large body size, especially for males who gain access to resources and mates^{11,24} (tall men included²⁵), it is unsurprising that many species of mammals, birds, fish, reptiles, amphibians, and arthropods have evolved anatomical or behavioural adaptations to exaggerate their apparent size^{6,19,26–28}. For example, the already-descended larynx of red deer stags²⁹ is lowered further still during roaring contests with rival males³⁰, extending the vocal tract even more to produce

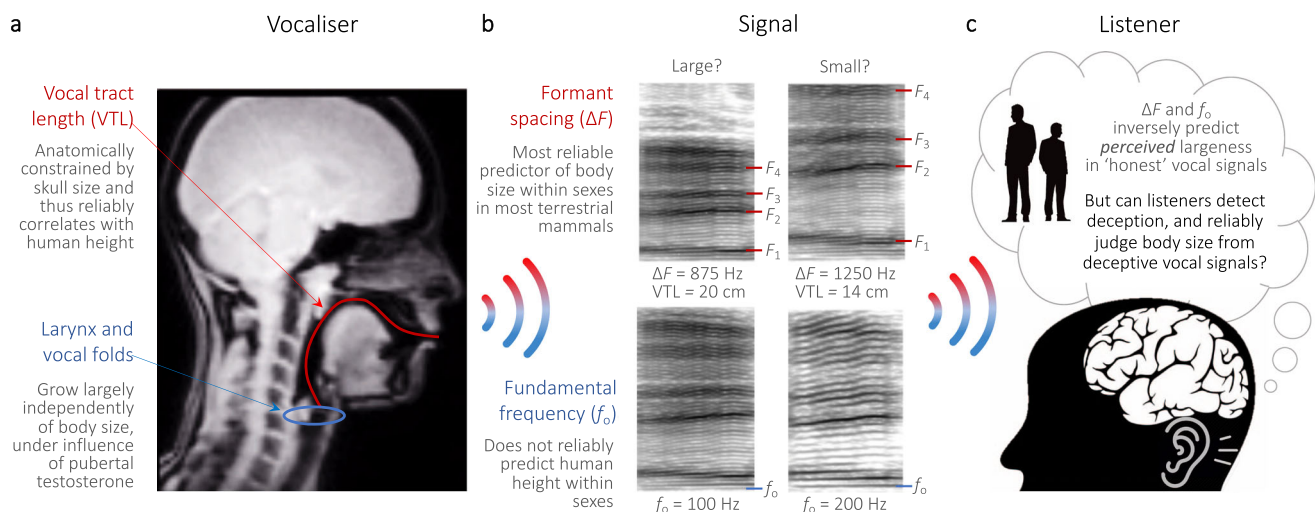


Fig. 1 Human vocal communication of body size. **a** Sagittal MRI image of human vocal anatomy illustrating vocal tract length (VTL) (in red) during production of the vowel /u:/, and permanently descended larynx housing the vocal folds (blue circle). Human VTL scales fairly allometrically with body size due to anatomical constraints, whereas vocal fold length does not^{26,28,32,33}. **b** Resonances of the vocal tract termed formants (labelled F_1 – F_4) and their relative spacing (ΔF) inversely predict human height controlling for sex and age^{33,34}, whereas fundamental frequency (labelled f_0), related to vocal fold length, does not³⁴. **c** Human listeners robustly associate both voice frequencies (low ΔF and f_0) with large body size in regular speech^{25,36} and in nonverbal vocalisations (e.g., roars³⁹), but can humans judge size from deceptive vocal signals? Panels **a**, **b** were reproduced and reformatted with permission²¹; parts of panel **c** were designed using resources from freepik.com and rawpixel.com (https://www.freepik.com/free-vector/illustration-business-people_2609966.htm#query=man%20black%20silhouette&position=17).

abnormally low formant frequency spacing (ΔF , the overall spacing between any two consecutive formants in the frequency domain, see Fig. 1) given the animal's true size^{27,31}.

Humans also possess a descended and sexually dimorphic larynx, with men boasting longer vocal tracts (reduced ΔF) and longer vocal folds (lower fundamental frequency or pitch, f_0) than women^{28,32,33} (Fig. 1). Although ΔF scales allometrically with vocal tract length (VTL) and thus predicts body size, both between and within adult sexes, f_0 is a poor predictor of human height at the intrasexual level^{28,33,34} (see Fig. 1). Yet, despite strongly associating not only ΔF but also f_0 with physical largeness^{25,35,36}, listeners can gauge relative body size from modal speech and nonverbal vocalisations^{25,36,37}. Critically, however, while we have recently shown that men and women can behaviourally lower their voice ΔF and f_0 to further exaggerate their body size and strength^{38,39}, remarkably little is known about the role of such deception in size communication.

In this study, we combine acoustic analysis of vocal signals, produced by men and women attempting to sound physically larger or smaller, with a series of psychoacoustic playback experiments conducted on a representative sample of 200 human listeners. Their task was to judge the heights of these vocalisers, and to attempt to discriminate among honest, exaggerated, or attenuated vocal signals of body size. Using this innovative approach we address long-standing questions about the evolution of deceptive signals in animal communication: Does deceptive size signalling retain an element of honesty? Can listeners detect size deception and do they correct for it when judging height? If so, does it still benefit vocalisers to exaggerate their perceived size? Are there sex differences, as predicted by sexual selection^{11,25}, in the production and perception of exaggerated signals? Our results offer a unique lens into the conflict between deceivers and receivers, showing that while deceptive vocal signals can effectively bias listeners' judgements of body size, such signals remain constrained and thus retain some reliable information. Specifically, we show that listeners often correctly discriminate between honest and deceptive vocal signals, and that when they do detect deception, they can recalibrate their height judgements accordingly. This research reveals that although listeners are not systematically fooled, it still pays off to deceive.

Results

Voice frequency shifts in size deception. As a first step toward answering our key research questions, we analysed speech signals (vowels /a i e o u/) produced by men and women tasked with sounding physically larger and smaller than their true body size³⁸. Acoustic analyses performed in PRAAT⁴⁰ (see Methods) confirmed that both sexes volitionally lowered their voice fundamental frequency (f_0) and formant spacing (ΔF) to deceptively exaggerate their apparent body size, and raised both frequency parameters to attenuate it, relative to their unmodulated (herein 'honest') vocal signals.

Men, by extending their apparent VTL more extremely than did women, lowered their formants more (Supplementary Tables 1 and 2), which is expected to simulate a larger body size (Fig. 1). In contrast, men did not raise their formants significantly more than did women to sound smaller, nor did they shift their voice pitch (f_0) more (LMM, Supplementary Table 2). The observed sex difference in formant shifts cannot be attributed to sexual dimorphism in starting VTL as controlling for baseline VTL or ΔF showed that the mean percentage change in men's formant shifts was still three to four times greater than that of women (Supplementary Table 1). Even within sexes, taller men likewise shifted their formants during size deception more than did shorter men, wherein relative differences in men's heights

explained one-third of the variance in formant shift magnitude for size exaggeration. Again, this relationship was not observed among women, nor for f_0 shifts in either sex (Supplementary Table 3). Taken together, these results support the hypothesis that the human male vocal tract may have evolved largely under selection pressure for size exaggeration^{21,28}, wherein men may perform vocal tract dynamics to maximise their perceived body size, and ultimately, their reproductive success²⁵, while simultaneously preserving phonetic range as illustrated by articulatory models⁴¹.

Formant measures (ΔF and apparent VTL) taken from central vowel frequencies in honest vocal signals explained between 14% and 40% of the variance in actual height within sexes, whereas f_0 explained virtually none (<3%; Supplementary Table 4). These findings corroborate studies on several other terrestrial mammals^{26,27} including humans (see the meta-analysis in ref. 34), where formants but not f_0 are anatomically constrained and thus follow a degree of acoustic allometry (Fig. 1). Importantly, formant measures also predicted inter-individual differences in height from deceptive vocal signals, particularly among men exaggerating their size (ΔF $R^2 = 0.58$) and women attenuating their size (ΔF $R^2 = 0.17$; Supplementary Table 4). Here too, f_0 did not significantly predict individual differences in actual height from deceptive signals ($R^2 = 0.003$ – 0.12 ; Supplementary Table 4). These acoustic analyses show that reliable formant-based information indicating inter-individual differences in body size is present in the human voice even during size deception, and thus, that listeners may be able to reliably gauge relative size from deceptive vocal signals.

Vocal size deception biases listeners. Do these deceptive signals fool listeners? To answer this imperative question, human adults (Experiment 1: $n = 97$, aged 18–63, 59 males) completed two psychoacoustic tasks involving (1) judging the absolute heights of vocalisers from their honest and deceptive vocal signals using a sliding metric/imperial scale, and (2) judging whether those same vocalisers were speaking naturally or instead deceptively exaggerating or attenuating their size (see Methods). To avoid cueing listeners to the possibility of vocal deception and thus to maintain ecological validity, listeners judged the height before assessing deception in a separate experimental block (but see Experiment 2).

Linear mixed models (LMMs) confirmed that listeners' height judgements were indeed biased by size deception when judging both male ($F_{2,1938} = 299.2$, $p < 0.001$) and female vocalisers ($F_{2,1938} = 174.3$ $p < 0.001$), with no effects of listener sex (Supplementary Table 5). As illustrated in Fig. 2a, on average, listeners overestimated the height of size exaggerators (estimated marginal means, M 3.4 cm, 95% CI 2.8, 4.1 male voices; M 1.9 cm, 95% CI 1.23, 2.6 female voices) and underestimated the height of size attenuators, by an average of 4 cm (M -4.3 cm, 95% CI -5.0, -3.7 male voices; M -4.0 cm, 95% CI -4.7, -3.3 female voices). Men attempting to exaggerate (but not attenuate) their size biased listeners' judgements more effectively than did women (Fig. 2a; $F_{2,3872} = 9.9$, $p < 0.001$; Supplementary Table 5), consistent with previous suggestions that the exaggeration of apparent body size is under stronger sexual selection in male than female vocal signals^{11,25,26,35}, especially in sexually dimorphic species in which males are larger than females¹⁹.

The moderate positive relationship between actual and perceived height in men's honest vocal signals ($r = 0.44$) was retained in their exaggerated ($r = 0.48$) but not attenuated signals (Fig. 2b). In women, honest ($r = 0.54$) and attenuated ($r = 0.50$) signals showed the strongest relationships (Fig. 2b). Thus, although listeners' individual height judgements were systematically shifted up by size

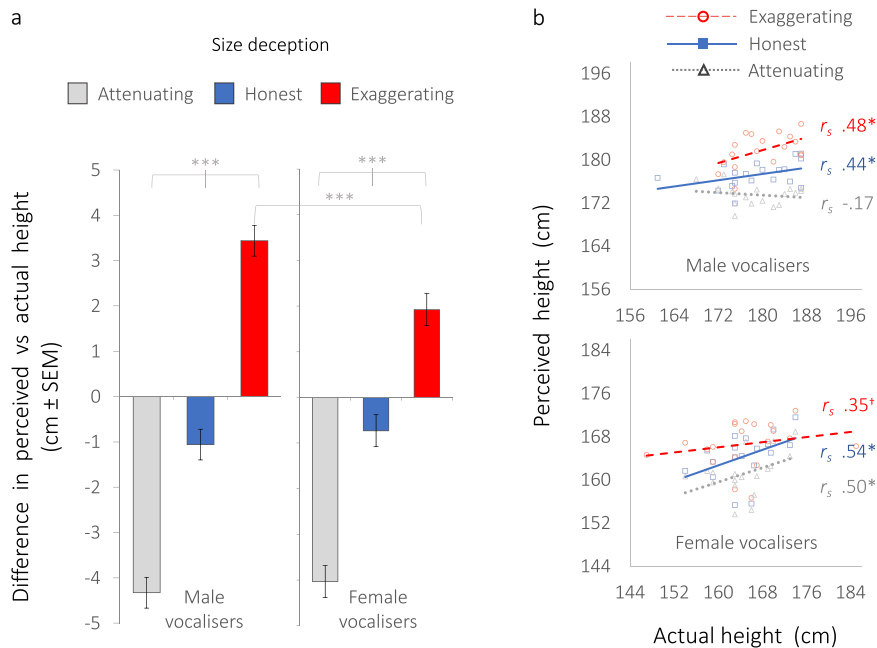


Fig. 2 Vocal size deception biases judgements of body size (Experiment 1). **a** Bias in height judgements shown as the mean difference (\pm SEM) between perceived and actual heights of vocalisers, in cm, for honest vocal signals (central blue bars) and deceptive vocal signals (attenuating size = grey bars, exaggerating size = red bars), where 0 indicates accurate height judgements, positive values indicate overestimation, and negative values indicate underestimation. Estimated marginal means and pairwise comparisons derive from linear mixed models, LMMs (see Supplementary Table 5), where all $***p < 0.001$ following Šidak correction for multiple comparisons. Tests are two-tailed. Error bars represent standard errors of the mean, \pm SEM. See also Supplementary Fig. 1 for dot plots illustrating the distribution of data. **b** Linear regressions showing relationships between perceived and actual heights of vocalisers, where each dot represents a vocaliser, and height judgements are averaged across listeners for each vocaliser and each size condition (exaggerating size = red circles, honest = blue squares, attenuating = grey triangles), where Cook's $D_i < 0.20$ (see Methods). 1 cm on the x-axis is equal to 1 cm on the y-axis. Spearman's ρ (r_s) correlation coefficients are given for each regression line, where $*p < 0.05$, one-tailed (males exaggerating $r = 0.48$, $p = 0.02$, $n = 18$; honest $r = 0.44$, $p = 0.03$, $n = 19$; attenuating $r = -0.17$, $p = 0.25$, $n = 18$; females exaggerating $r = 0.35$, $p = 0.06$, $n = 20$; honest $r = 0.54$, $p = 0.01$; $n = 18$; attenuating $r = 0.50$, $p = 0.02$; $n = 18$). All data were derived from Experiment 1 based on 120 vocal stimuli produced by $n = 40$ vocalisers (20 males, 20 females) in each of three size conditions (honest, attenuating, exaggerating) and judged by $n = 97$ listeners, where each vocal stimulus was rated by an average of 50 listeners (see Methods). Source data are provided as a Source Data file.

exaggeration and down by size attenuation, inter-individual differences in height were preserved in listeners' judgements. This could be expected based on our acoustic analyses indicating reliable formant-based cues to actual size in deceptive vocal signals (Supplementary Table 4). As further predicted, listeners generally associated lower absolute voice frequencies with larger size, a well-established perceptual association^{25,27,36,37}. However, this association was stronger for exaggerated than honest or attenuated signals (Supplementary Table 6). Together these results suggest that lowering voice frequencies can function to effectively maximise perceived absolute body size, again to the potential fitness benefit of the size exaggerator^{11,25,35}.

Listeners can detect deception. In a second task, the same listeners judging the same vocal stimuli correctly detected the presence or absence of size deception above chance (Fig. 3a; chance = 33%, grand mean correct $51\% \pm 0.007$ SEM, 95% CI 49–52%) and more reliably detected deception by men than deception by women during intended size exaggeration ($51\% \pm 0.02$ vs $45\% \pm 0.02$) and attenuation ($49\% \pm 0.02$ vs $40\% \pm 0.02$) with no differences between male and female listeners (LMM, Supplementary Table 7, Fig. 3a).

A crucial question then is whether listeners recalibrated their height judgements for voice signals that they were able to correctly detect as deceptive in a subsequent task. As illustrated in Fig. 3b (LMMs, Supplementary Tables 8 and 9), honest signals correctly identified as such were indeed associated with reliable

height judgements overall (i.e., error near 0, $M -0.035$ cm, 95% CI $-0.8, -0.78$ male voices; $M 0.009$ cm, 95% CI $-0.87, 0.89$ female voices), whereas honest signals misidentified as deceptive were underestimated ($M -2.6$ cm, 95% CI $-3.6, -1.6$ male voices; $M -1.7$ cm, 95% CI $-2.7, -0.7$ female voices). Surprisingly, however, listeners' judgements of height were maximally biased for deceptive signals they correctly identified as deceptive in a subsequent task (Fig. 3b, c).

These results indicate that listeners remain partly 'fooled' by vocal deception, and were in fact most fooled by voices they could correctly identify as deceptive in a separate task. We examined the possibility that detectable deceivers shifted their voice frequencies more than those who remained undetected, which could increase detection while also maximally exploiting deep-seated sound-size correspondences between low frequencies and largeness^{25,27,36,37}. Although women who were detected as exaggerating did indeed shift their ΔF and f_0 more than those who remained undetected (Supplementary Table 10), such acoustic differences were not observed among men nor in the context of size attenuation. Hence, we tested a second possibility that the listeners in this experiment, who were not 'primed' to contemplate deception when judging the heights of vocalisers, may have failed to detect deception at that stage, and thus to recalibrate their judgements.

Awareness of deception reduces bias. To test this possibility, we conducted a second psychoacoustic experiment on an independent

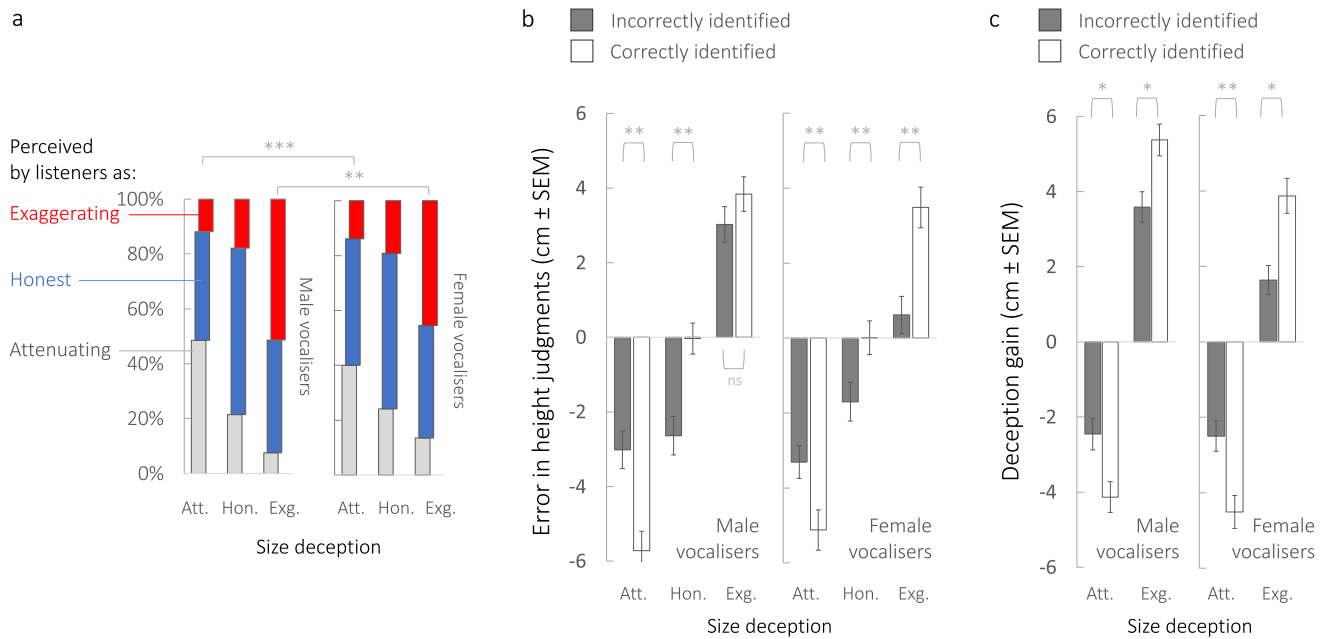


Fig. 3 Listeners can detect deception, but remain deceived by deceptive signals (Experiment 1). **a** Percentages of vocalisers that listeners perceived as deceptively exaggerating (red bars) or attenuating (light grey bars) their size, or as producing honest vocal signals (blue bars, centre) are shown along the y-axis as a function of the intended size deception indicated along the x-axis. Estimated marginal means and pairwise comparisons derive from LMMs (see Supplementary Table 7), where $***p < 0.001$, $**p < 0.01$, following Šidák correction for multiple comparisons. Tests are two-tailed. **b, c** Bias in listeners' size assessments as a function of whether a listener failed to detect (dark grey bars) or correctly detected (white bars) a vocal signal as deceptive or honest, where panel **b** shows 'error' in height judgements (mean difference between perceived vs actual heights of vocalisers), and panel **c** shows 'deception gain' in height judgements (mean difference between perceived height from honest signals and perceived height from deceptive signals). Estimated marginal means and pairwise comparisons derive from LMMs (see Supplementary Tables 8 and 9), $**p < 0.01$, $*p < 0.05$ following Šidák correction. Tests are two-tailed. Error bars \pm SEM. Acronyms: Att. attenuating, Hon. honest, Exg. exaggerating. All data derive from Experiment 1, based on 120 vocal stimuli produced by $n = 40$ vocalisers (20 males, 20 females) in each of three size conditions (honest, attenuating, exaggerating) and judged by $n = 97$ listeners, where each vocal stimulus was rated by an average of 50 listeners (see Methods). See also Supplementary Fig. 2 for dot plots illustrating distribution of data. Source data are provided as a Source Data file.

sample of listeners (Experiment 2: $n = 98$, aged 18–71, 59 males). The experiment was identical to the first, except that listeners now completed both tasks concurrently, first indicating whether or not they perceived a vocaliser as deceptively altering their size, and then judging the height of that same vocaliser within the same trial (see Methods). The results confirmed our prediction (Fig. 4; LMMs, Supplementary Tables 11–14). Indeed, listeners primed to seek deceit were substantially less biased by vocal size exaggeration, overestimating or underestimating the heights of vocal deceivers by half the magnitude as listeners in Experiment 1 (Fig. 4a vs Fig. 2a and Supplementary Table 11). Although listeners in both experiments detected size deception with similar verity (Fig. 4b vs Fig. 3a and Supplementary Table 12), critically, listeners in Experiment 2 more effectively recalibrated their height judgements for vocal signals they correctly identified as deceptive (Fig. 4c, d vs Fig. 3b, c and Supplementary Tables 13, 14). This effect was most pronounced for size exaggerators who, when correctly detected as cheating, failed to fool listeners into perceiving them as much larger than their true body size (Fig. 4c, d). Moreover, male listeners recalibrated their size judgements after correctly detecting deception more effectively than did female listeners, specifically when assessing the body size of other men (Fig. 4d left panel and LMM, Supplementary Table 14).

Discussion

Our results provide rare insight into an evolutionary arms race between signallers and receivers, addressing several long-standing questions aimed at resolving the honest signalling paradox^{1–9}. First, is honesty retained in deceptive signals? Yes. Presumably

due to anatomical constraints on VTL^{26,28} limiting the extent of deception, the relationship between formant frequencies and true body size (acoustic allometry) is preserved, albeit shifted, during vocal size deception. The resulting reliable cues to intra-individual differences in height ostensibly allowed listeners in the present study to 'rank' the relative heights of cheating vocalisers. Second, can listeners detect deception? Yes. Listeners correctly identified vocal signals as honest or deceptive much more often than expected by chance, though they nevertheless failed to detect deception approximately half of the time. Third, do listeners recalibrate their height judgements when they detect deception? Yes, largely, and much more if they are primed to seek deception. Indeed in Experiment 2, height judgements were significantly less biased for vocal signals correctly and concurrently detected as deceptive. Fourth, and crucially, does it still pay to deceive? Yes. We show that vocalisers who attempted to alter their apparent body size by shifting the frequency parameters of their voice effectively fooled listeners, who indeed perceived them as taller or shorter, by several centimetres on average. Although primed detection of deception reduced this bias, listeners' height judgements remained biased overall, as deceit often went undetected.

Our results also show that males more effectively exaggerate their body size through voice modulation than do females. Despite being detected as cheating more often than women, we show that men, who stand to gain relatively greater fitness benefits by successfully exaggerating their size^{11,25,35}, shifted their formant frequencies more than did women. Men thus simulated a longer vocal tract and larger body size, and more effectively biased listeners to overestimate their height. From the receiver's

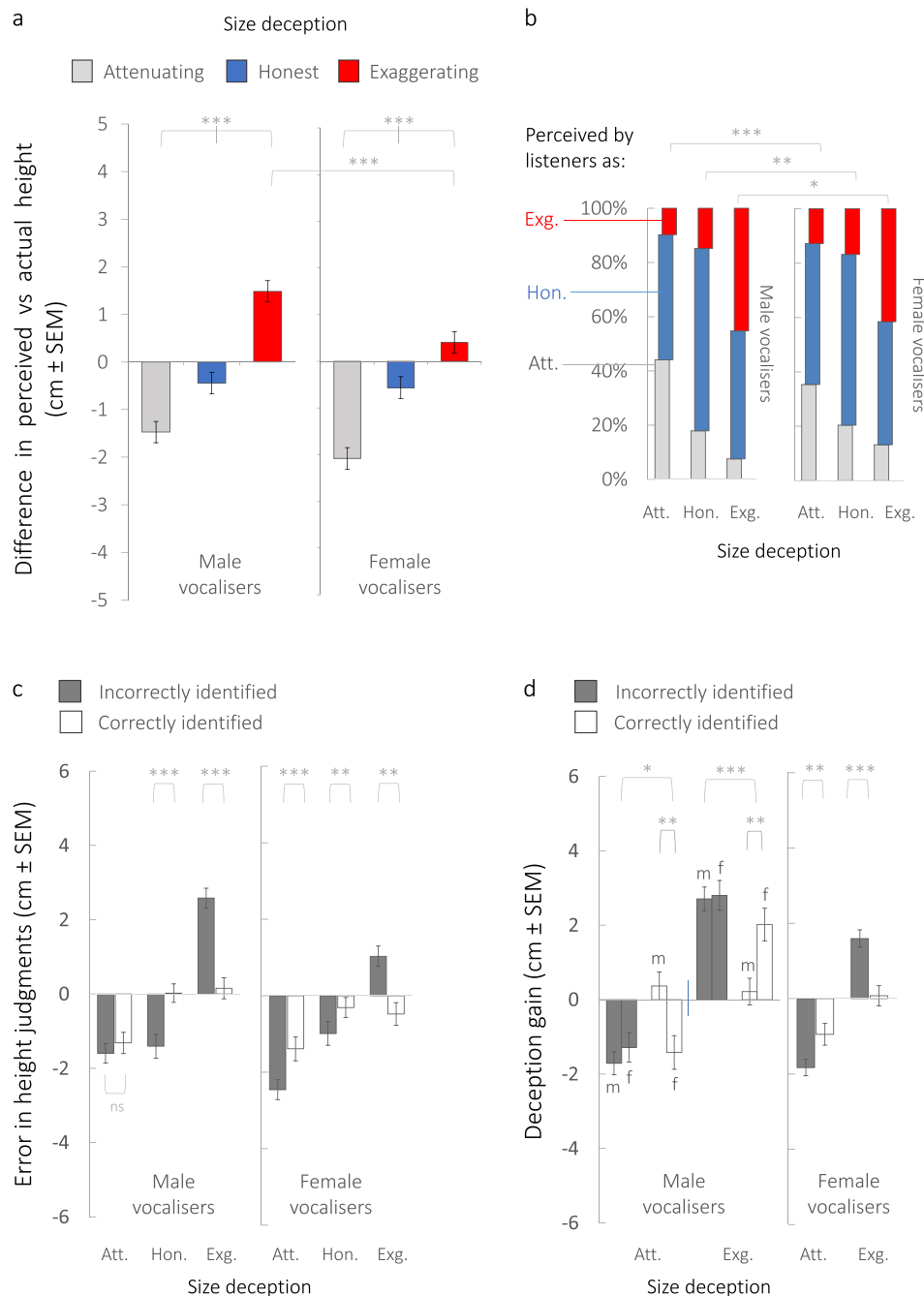


Fig. 4 Awareness reduces bias: Listeners recalibrate height judgements for signals correctly and concurrently detected as deceptive (Experiment 2).

a Bias in height judgements shown as the mean difference (\pm SEM) between perceived and actual heights of vocalisers, in cm, for honest vocal signals (blue bars) and deceptive vocal signals (attenuating = grey bars, exaggerating = red bars). Estimated marginal means and pairwise comparisons derive from LMMs (see Supplementary Table 11), where all $***p < 0.001$ following Šidák correction for multiple comparisons. Error bars \pm SEM. **b** Percentages of vocalisers that listeners perceived as deceptively exaggerating (red bars) or attenuating (light grey bars) their size, or as producing honest vocal signals (blue bars, centre) are shown along the y-axis as a function of the intended size deception indicated along the x-axis. Estimated marginal means and pairwise comparisons derive from LMMs (see Supplementary Table 12), where $***p < 0.001$, $**p < 0.01$, $*p < 0.05$ following Šidák correction. Tests are two-tailed. **c, d** Bias in listeners' size assessments as a function of whether a listener failed to detect (dark grey bars) or correctly detected (white bars) a vocal signal as deceptive or honest, where panel **c** shows 'error' in height judgements (mean difference between perceived and actual heights of vocalisers), and panel **d** shows 'deception gain' in height judgements (mean difference between perceived height from honest signals and perceived height from deceptive signals). Panel **d** also illustrates that deception gain was lower for male vocalisers (left side) when correctly detected as cheating by other male listeners (labelled with 'm') compared to when detected by female listeners ('f'). Estimated marginal means and pairwise comparisons derive from LMMs (see Supplementary Tables 13 and 14), $***p < 0.001$, $**p < 0.01$, $*p < 0.05$ following Šidák correction. Tests are two-tailed. Error bars \pm SEM. Acronyms: Att. attenuating, Hon. honest, Exg. exaggerating, m male listeners, f female listeners. All data derive from Experiment 2 based on 120 vocal stimuli produced by $n = 40$ vocalisers (20 males, 20 females) in each of three size conditions (honest, attenuating, exaggerating) and judged by $n = 98$ listeners, each of whom rated all 120 vocal stimuli (see Methods). See also Supplementary Fig. 3 for dot plots illustrating the distribution of data. Source data are provided as a Source Data file.

side, we also found that, when primed to the possibility of deception, male listeners were less susceptible than female listeners to deceptive signals produced by other men if they correctly detected them as exaggerating or attenuating their body size. The specificity of this effect to the second experiment suggests that male listeners are particularly attuned to the deceptive signals of other men when an explicit competitive context is induced. Together these findings support the prediction that pressure on listeners to counteract deception by recalibrating size judgements for deceptive signals may be maximised in the context of male–male competition. Finally, our finding that listeners can effectively gauge the relative heights of deceivers predicts an asymmetry in the impact of deception on male and female listeners⁶. Indeed, assuming all males exaggerate, females should retain the ability to rank relative male quality, which is crucial in mate choice. In contrast, males may overestimate the absolute size of exaggerating competitors when deception goes undetected. In male–male competition, the size and strength of a rival compared to oneself is critical³⁷, and thus males may overvalue the cost of continued conflict. This sexual asymmetry is again consistent with the dominant view that male–male competition is the primary mechanism of selection on men’s sexually dimorphic traits⁴², including voice pitch⁴³, and here, a probable key driver of size exaggeration.

Honesty has been identified in exaggerated vocal signals of other species, such as red deer^{29,30,44} and koalas^{45,46} who have both permanent and behavioural adaptations for size exaggeration and where physical constraints enforce reliability. Hypothetically this may lead receivers to adapt to deception by shifting their perceptual scale to exaggerated ranges. Here, we show that absolute deception remains effective, indicating that such perceptual shifts are only partial. We suggest that similar effectiveness may be predicted in nonhuman signalling systems where deception is behavioural and facultative, and especially where constraints enforce a degree of reliability in the signalling of interindividual differences.

Body size exaggeration is suspected in numerous species across a range of taxa^{6,8,19,26} however, to our knowledge, its effectiveness in biasing listeners had never been established. While researchers often generalize animal models to humans, our work shows that studying the human animal can answer key questions about animal behaviour that are otherwise difficult to tackle (see also⁴⁷). At the same time, humans are exceptionally complex. Beyond body size, human voice fundamental and formant frequencies predict a range of biologically and socially relevant traits, such as dominance, strength, and attractiveness^{25,35,48}, that can interact cross-modally with visual and olfactory cues to influence listeners’ perceptions of the signaler. Voice modulation can also communicate motivation (e.g., aggressive intent) rather than physical traits per se^{49,50}. Research into vocal deception of a wide range of traits⁵¹ and states⁵⁰, particularly in multimodal real-world contexts⁵², is needed to further elucidate the functions and tangible consequences of deceit in complex social environments. Humans also possess an unprecedented capacity for volitional vocal control²¹ and an advanced theory of mind. While intentionality is not necessary for deceptive signalling to evolve (i.e., functional deception⁵), this has nevertheless led many researchers to suggest that ‘human communication is permeated with deceit’⁶, perhaps more so than the communication systems of other animals. Indeed, our results suggest that priming listeners to deception can substantially decrease its effectiveness and thus reduce potential costs of being deceived for cognisant listeners. Further comparative work is needed to elucidate the cognition of deception⁸, in humans and nonhuman animals.

Methods

Vocal stimuli. Vocal stimuli derived from 40 adult English speakers: 20 men (mean age 19.6 ± 2.4 sd) whose heights ranged from 161 cm to 187 cm (mean height 178.4 ± 7 cm sd) and 20 women (mean age 19.1 ± 1.6 sd) whose heights ranged from 147 to 185 cm (mean height 164.9 ± 7.9 sd), taken from a larger sample of vocal stimuli (see³⁸). The stimuli were selected to be representative of a broad range of heights: the height distributions of male and female vocalisers closely parallel those observed in large cross-cultural samples of adults (men 178 ± 6.58 cm, $n = 1334$; women 165.96 ± 6.64 cm, $n = 871$)³⁴. Vocalisers were recorded in an anechoic sound-controlled chamber with a Sennheiser MKH 800 cardioid condenser microphone at a distance of approximately 10 cm. Voice stimuli consisted of a series of monophthong vowels (/a/, /i/, /e/, /o/, /u/). Each vocaliser produced the vowels in three conditions, beginning with a baseline ‘honest’ condition in which they spoke the vowels in their natural voice. They were then asked to reproduce the vowels while sounding physically large, and again while sounding physically small, in a counter-balanced order across participants³⁸. No further instructions were given. Recordings were digitally encoded with an M-Audio Fast Track interface at a sampling rate of 96 kHz and 24-bit amplitude quantisation, and stored onto a computer as PCM WAV files. This procedure resulted in 120 vocal stimuli.

Acoustic analysis. Fundamental frequency (f_0) and the first four formant frequencies (F_1 – F_4) were measured using the well-established autocorrelation algorithm (f_0 range 30–500 Hz for men and 65–600 Hz for women) and Burg linear predictive coding algorithm (max formant 5000 Hz for men and 5500 Hz for women) in PRAAT acoustic software⁴⁰. Formant measures were taken from the mean centre frequencies of each vowel and averaged within vocalisers and voice conditions. From F_1 to F_4 we computed formant spacing (ΔF), a measure of the distance among adjacent formants, and apparent VTL^{31,34}, an estimate of the length of the supralaryngeal vocal tract (Fig. 1), both of which explain the highest proportion of variance in height among humans³⁴ and many other mammals²⁷ within sex-age classes. Mean f_0 measures were additionally transformed into equivalent rectangular bandwidth units (ERB, where $E_i = 21.4^* \log_{10}(0.00437^* f_i + 1)$)⁵³ and formant measures were transformed into Bark units (where $Z_i = 26.81/(1 + 1960/f_i) - 0.53$)⁵⁴. These quasi-logarithmic scales control for any difference between the physical and perceived properties of these frequencies. However, due to extremely strong collinearity between the two measures of mean f_0 (Hz and ERB, $r = 0.99$) and among various formant measures (ΔF and VTL in Hz and Bark, $r = -0.99$), we found virtually identical results regardless of which measure was used for each given vocal parameter, and thus, data are presented for ΔF and f_0 in Hz only.

Listeners. Two-hundred adult listeners took part in one of two psychoacoustic experiments. Sample sizes for Experiment 1 were determined prior to testing to achieve an average of 50 height judgements per vocal stimulus for a statistical power of 80%, in order to obtain a small-to-medium effect size in regressions between perceived and actual vocaliser height. While high inter-rated agreement (alphas > 0.80 , $p_s < 0.001$) among listeners is typically achieved with relatively small sample sizes (e.g., less than 15 listeners per sex for voice-based judgements of dominance or attractiveness⁴³), earlier studies on human vocal communication of body size have generally failed to find significant correlations between perceived and actual height in one or both sexes of vocalisers with samples of fewer than 25 listeners per vocal stimulus^{55–57}. In Experiment 2, tasks were conjoined, and thus participants rated all vocal stimuli (see Psychoacoustic playback experiments). English-speaking listeners were recruited via Amazon’s online recruitment platform, Mechanical Turk. The use of headphones was mandatory. In Experiment 1, three participants did not finish the study and were thus excluded from analyses. In the final sample ($n = 97$), 59 participants indicated male gender (aged 18–63, 9.2 sd), 36 indicated female gender (aged 18–63, 11.1 sd), and two indicated their gender as ‘other’. In Experiment 2, two participants provided random responses and were thus excluded from analyses. In the final sample ($n = 98$), 59 participants indicated male gender (aged 21–71, 9.7 sd) and 39 indicated female gender (aged 18–55, 9.4 sd). This research was approved by the Institutional Ethics Committee (C-REC; Certificates of approval: ER/KP292/11 and ER/REBY/12). All participants provided informed consent and were reimbursed monetarily at the recommended rate of \$0.13 USD per minute⁵⁸.

Psychoacoustic playback experiments. Listeners completed a short demographic questionnaire and took part in one of the two psychoacoustic playback experiments, custom designed in Syntoolkit⁵⁹, each involving two tasks: judging vocaliser height and detecting vocal size deception. In Experiment 1, listeners performed these two tasks in separate, consecutive blocks. In each task listeners rated the voices of a random sample of 10 male and 10 female vocalisers, in each of three voice conditions (honest, exaggerated, and attenuated size), resulting in 60 trials per task, or a total of 120 trials per listener. Height judgements preceded the deception detection task so as not to prime nor bias listeners toward contemplating deception when judging height, thus reflecting a more ecologically valid experimental design. Importantly, the same voice stimuli were presented in both tasks within listeners to allow for meaningful comparisons. In Experiment 2, an independent sample of listeners performed the same tasks; however, the tasks were now performed concurrently for each vocal stimulus. Thus, listeners first indicated

whether or not they perceived a vocaliser as deceptively altering their voice to sound larger or smaller, and then judged the height of that same vocaliser, within the same experimental trial. Listeners judged all 20 male and 20 female vocalisers, in each of three voice conditions (honest, exaggerated, and attenuated size), for a total of 120 trials per listener. In both Experiments, listeners were presented with a single vocal stimulus on each trial. Voice stimuli were blocked by the sex of the vocaliser, and block order and stimulus presentation within each block were randomized. Listeners were instructed to wear headphones and not to adjust their volume settings throughout the experiment; this was verified during debriefing.

Task 1: Judging height. To indicate the perceived height of a vocaliser, listeners used a vertical sliding bar, which appeared only after a voice stimulus finished playing. As the participant moved the cursor along the vertical sliding bar, the selected height was indicated in both metric (cm) and imperial (feet and inches) units. Maximum and minimum heights were labelled on the top and bottom of the scale, respectively, based on sex-specific distributions of heights in our samples, which correspond closely with those observed in the general population³⁴. Thus, the centre of the sliding bar was set to 178 cm for men and 165 cm for women, with end points corresponding to three standard deviations above and below these means: a range of 156–198 cm for men and 144–186 cm for women.

Task 2: Detecting deception. To indicate perceived vocal size deception or its absence, listeners were instructed to indicate whether they believed the vocaliser was speaking in their natural voice or changing their voice to sound physically smaller or larger than they actually are. These three options, presented as radial buttons, appeared only after the voice stimulus finished playing.

Statistical analysis. A series of LMMs fit by restricted maximum likelihood estimation were used to examine listeners' height judgements, correct detection of vocal size deception, and the influence of this detectable deception on height judgements. The key variable of interest (size deception: honest, attenuated, exaggerated) was entered in LMMs as a fixed variable. Vocaliser and listener IDs were always included as random variables with random intercepts in all models, and the sex of both vocaliser and listener were entered as fixed variables in omnibus models. Vocaliser sex consistently showed significant effects; therefore separate LMMs are reported for male and female vocalisers. Where there were no significant effects of listener sex (Experiment 1) listener data were pooled for all analyses; otherwise listener sex was retained in final models where applicable. Full model parameters of LMMs are detailed below each respective output table (see Supplementary Tables 2, 5, 7, 8, 9, 11, 12, 13 and 14). All means, confidence intervals (95% CI), and standard errors (SEM) reported in the paper derive from LMMs. Significant effects in LMMs were further examined using pairwise tests with Sidák correction for multiple comparisons. Multiple and simple linear regression models were employed to examine relationships among continuous variables. Where applicable, height judgements were averaged across listeners for each vocaliser and voice condition. Spearman's rho (r_s) correlation coefficients were used where data were nonnormally distributed or potentially nonlinear, and these statistics are reported in the paper, however Pearson's r coefficients are given for comparison in Supplementary Tables 3, 4, and 6. Cook's distances were calculated to identify influential outliers in simple regressions (where $D_i > 0.20$), and the statistics reported in the paper exclude outliers. Removing or retaining outliers did not affect the direction and general pattern of relationships. We used an alpha level of 0.05 for all statistical tests, and all tests were two-tailed with the exception of simple regressions, where we had a priori directional predictions. Error bars in all figures represent standard errors of the mean (\pm SEM), whereas 95% CI are given in text. Statistical analyses were performed in SPSS 24 (IBM). Full statistical models and data are available in the supplementary files of this article (see Supplementary Data 1 and Supplementary Information).

Ethics statement. This research was approved by the University of Sussex's Life Sciences & Psychology Cluster-based Research Ethics Committee (C-REC; Certificates of approval: ER/KP292/11 and ER/REBY/12) and complies with the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct, including obtaining informed consent from all human participants.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files, including source data for all figures provided as a Source Data file. Datasets and voice stimuli ($n = 120$ audio WAV files) are also available on the Open Science Framework (<https://osf.io/r7gzb/>, <https://doi.org/10.17605/OSF.IO/R7GZB>). Source data are provided with this paper.

Received: 14 September 2020; Accepted: 5 January 2021;

Published online: 12 February 2021

References

- Zahavi, A. Mate selection—a selection for a handicap. *J. Theor. Biol.* **53**, 205–214 (1975).
- Dawkins, R. & Krebs, J. R. Animal signals: information or manipulation. *Behav. Ecol. Evol. Approach* **2**, 282–309 (1978).
- Maynard-Smith, J. & Parker, G. A. The logic of asymmetric contests. *Anim. Behav.* **24**, 159–175 (1976).
- Johnstone, R. A. & Grafen, A. Dishonesty and the handicap principle. *Anim. Behav.* **46**, 759–764 (1993).
- Hauser, M. D. *The Evolution of Communication* (MIT Press, 1996).
- Searcy, W. A. & Nowicki, S. *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems* (Princeton Univ. Press, 2005).
- Carazo, P. & Font, E. 'Communication breakdown': the evolution of signal unreliability and deception. *Anim. Behav.* **87**, 17–22 (2014).
- Mitchell, R. W. & Thompson, N. S. (eds.). *Deception: Perspectives on Human and Nonhuman Deceit* (SUNY Press, 1986).
- Reid, S. A., Zhang, J., Anderson, G. L. & Keblusek, L. Costly signaling in human communication. In *The Handbook of Communication Science and Biology* (eds. Floyd, K. & Weber, R.) (Routledge, 2020).
- Rendall, D., Owren, M. J. & Ryan, M. J. What do animal signals mean? *Anim. Behav.* **78**, 233–240 (2009).
- Andersson, M. B. *Sexual Selection* (Princeton Univ. Press, 1994).
- Flower, T. P., Gribble, M. & Ridley, A. R. Deception by flexible alarm mimicry in an African bird. *Science* **344**, 513–516 (2014).
- Hurd, P. L. Is signalling of fighting ability costlier for weaker individuals? *J. Theor. Biol.* **184**, 83–88 (1997).
- Maynard-Smith, J. & Price, G. R. The logic of animal conflict. *Nature* **246**, 15–18 (1973).
- Maynard-Smith, J. & Harper, D. *Animal Signals* (Oxford Univ. Press, 2003).
- Schaefer, M. & Ruxton, G. By-product information can stabilize the reliability of communication. *J. Evol. Biol.* **25**, 2412–2421 (2012).
- Folstad, I. & Karter, A. J. Parasites, bright males, and the immunocompetence handicap. *Am. Nat.* **139**, 603–622 (1992).
- Silk, J. B., Kaldor, E. & Boyd, R. Cheap talk when interests conflict. *Anim. Behav.* **59**, 423–432 (2000).
- Charlton, B. D. & Reby, D. The evolution of acoustic size exaggeration in terrestrial mammals. *Nat. Commun.* **7**, 12739 (2016).
- Arnott, G. & Elwood, R. W. Signal residuals and hermit crab displays: flaunt it if you have it! *Anim. Behav.* **79**, 137–143 (2010).
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J. & Reby, D. Voice modulation: a window into the origins of human vocal control? *Trends Cogn. Sci.* **20**, 304–318 (2016).
- Ackermann, H., Hage, S. R. & Ziegler, W. Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* **37**, 529–546 (2014).
- Belyk, M. & Brown, S. The origins of the vocal brain in humans. *Neurosci. Biobehav. Rev.* **77**, 177–193 (2017).
- Grafen, A. Biological signals as handicaps. *J. Theor. Biol.* **144**, 517–546 (1990).
- Pisanski, K. & Bryant, G. A. The evolution of voice perception. In *The Oxford Handbook of Voice Studies* (eds. Eidsheim, N. S. & Meizel, K. L.) (Oxford Univ. Press, 2019).
- Fitch, W. T. & Hauser, M. D. Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals. In *Acoustic Communication* (eds. Simmons, A. M., Fay, R. R. & Popper, A. N.) 65–137 (Springer, 2003).
- Charlton, B. D., Pisanski, K., Raine, J. & Reby, D. Coding of static information in terrestrial mammal vocal signals. In *Animal Signals and Communication* (eds. Aubin, T. & Mathevon, N.) 115–136 (Springer Nature, 2020).
- Fitch, W. T. The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267 (2000).
- Fitch, W. T. & Reby, D. The descended larynx is not uniquely human. *Proc. Biol. Sci.* **268**, 1669–1675 (2001).
- Reby, D. et al. Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proc. Biol. Sci.* **272**, 941–947 (2005).
- Reby, D. & McComb, K. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Anim. Behav.* **65**, 519–530 (2003).
- Titze, I. R. *Principles of Vocal Production* (Prentice-Hall, 1994).
- Fitch, W. T. & Giedd, J. Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoust. Soc. Am.* **106**, 1511–1522 (1999).
- Pisanski, K. et al. Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* **95**, 89–99 (2014).
- Putts, D. Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proc. Biol. Sci.* **283**, 20152830 (2016).
- Pisanski, K. & Rendall, D. The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *J. Acoust. Soc. Am.* **129**, 2201 (2011).
- Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J. & Reby, D. Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience* **4**, 273–280 (2018).

38. Pisanski, K. et al. Volitional exaggeration of body size through fundamental and formant frequency modulation in humans. *Sci. Rep.* **6**, 34389 (2016).
39. Raine, J., Pisanski, K., Bond, R., Simner, J. & Reby, D. Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech. *PLoS One* **14**, e0213034 (2019).
40. Boersma, P. & Weenink, D. Praat: Doing phonetics by computer v 6.1.21. <https://www.fon.hum.uva.nl/praat/> (2020).
41. De Boer, B. Investigating the acoustic effect of the descended larynx with articulatory models. *J. Phon.* **38**, 679–686 (2010).
42. Puts, D. A. Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* **31**, 157–175 (2010).
43. Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J. & Penke, L. The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evol. Hum. Behav.* **39**, 424–436 (2018).
44. Charlton, B. D., Reby, D. & McComb, K. Effect of combined source (F0) and filter (formant) variation on red deer hind responses to male roars. *J. Acoust. Soc. Am.* **123**, 2936 (2008).
45. Charlton, B. D. et al. Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Curr. Biol.* **23**, R1035–R1036 (2013).
46. Charlton, B. D. et al. Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: honesty in an exaggerated trait. *J. Exp. Biol.* **214**, 3414–3422 (2011).
47. Wilson, M. L., Miller, C. M. & Crouse, K. N. Humans as a model species for sexual selection research. *Proc. Biol. Sci.* **284**, 20171320 (2017).
48. Schild, C. et al. Linking human male vocal parameters to perceptions, body morphology, strength and hormonal profiles in contexts of sexual selection. *Sci. Rep.* **10**, 1–16 (2020).
49. Morton, E. S. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869 (1977).
50. Zhang, J., Hodges-Simeon, C., Gaulin, S. J. & Reid, S. A. Pitch lowering enhances men's perceived aggressive intent, not fighting ability. *Evol. Hum. Behav.* **42**, 51–60 (2020).
51. Hughes, S. M., Mogilski, J. K. & Harrison, M. A. The perception and parameters of intentional voice manipulation. *J. Nonverbal Behav.* **38**, 107–127 (2014).
52. Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M. & Reby, D. Voice pitch modulation in human mate choice. *Proc. Biol. Sci.* **285**, 20181634 (2018).
53. Glasberg, B. R. & Moore, B. C. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* **47**, 103–138 (1990).
54. Traunmüller, H. Auditory scales of frequency representation. (1997).
55. van Dommelen, W. A. & Moxness, B. H. Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Lang. Speech* **38**, 267–287 (1995).
56. Collins, S. A. Men's voices and women's choices. *Anim. Behav.* **60**, 773–780 (2000).
57. Bruckert, L., Lienard, J.-S., Lacroix, A., Kreutzer, M. & Leboucher, G. Women use voice parameters to assess men's characteristics. *Proc. R. Soc. B Biol. Sci.* **273**, 83–89 (2006).
58. Chandler, J. & Shapiro, D. Conducting clinical research using crowdsourced convenience samples. *Annu. Rev. Clin. Psychol.* **12**, 53–81 (2016).
59. Hughes, J. E., Gruffydd, E., Simner, J. & Ward, J. Synaesthetes show advantages in savant skill acquisition: training calendar calculation in sequence-space synaesthesia. *Cortex* **113**, 67–82 (2019).

Acknowledgements

This research and both authors K.P. and D.R. were funded by the University of Lyon IDEX Lyon 'Programme Investissements d'Avenir' (ANR-16-IDEX-0005) to D.R. and supported by the Labex CeLyA. Author K.P. was also supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 655859. Publication of this article in open access was financially supported by the Excellence Initiative - Research University (IDUB) programme for the University of Wrocław. The authors thank Profs M. Greenfield, M. Hauber and N. Mathevon for their constructive feedback on earlier versions of this paper.

Author contributions

K.P. and D.R. conceived and designed the study; K.P. collected all data, conducted analyses, and created all figures and tables; K.P. and D.R. interpreted the results and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21008-7>.

Correspondence and requests for materials should be addressed to K.P.

Peer review information *Nature Communications* thanks W. Tecumseh Fitch, Nate Pipitone and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021