

SCIENTIFIC REPORTS



OPEN

Volitional exaggeration of body size through fundamental and formant frequency modulation in humans

Katarzyna Pisanski^{1,2,3}, Emanuel C. Mora^{4,5}, Annette Pisanski⁴, David Reby³, Piotr Sorokowski², Tomasz Frackowiak² & David R. Feinberg¹

Received: 15 February 2016
Accepted: 09 September 2016
Published: 30 September 2016

Several mammalian species scale their voice fundamental frequency (F_0) and formant frequencies in competitive and mating contexts, reducing vocal tract and laryngeal allometry thereby exaggerating apparent body size. Although humans' rare capacity to volitionally modulate these same frequencies is thought to subserve articulated speech, the potential function of voice frequency modulation in human nonverbal communication remains largely unexplored. Here, the voices of 167 men and women from Canada, Cuba, and Poland were recorded in a baseline condition and while volitionally imitating a physically small and large body size. Modulation of F_0 , formant spacing (ΔF), and apparent vocal tract length (VTL) were measured using Praat. Our results indicate that men and women spontaneously and systemically increased VTL and decreased F_0 to imitate a large body size, and reduced VTL and increased F_0 to imitate small size. These voice modulations did not differ substantially across cultures, indicating potentially universal sound-size correspondences or anatomical and biomechanical constraints on voice modulation. In each culture, men generally modulated their voices (particularly formants) more than did women. This latter finding could help to explain sexual dimorphism in F_0 and formants that is currently unaccounted for by sexual dimorphism in human vocal anatomy and body size.

Several mammalian species are known to scale their vocal frequencies (see refs 1–4 for reviews). Red deer (*Cervus elaphus*) offer a prime example, wherein stags drastically lower their larynges to extend their vocal tracts during roaring and do so predominantly in response to threatening male competitors^{5,6}. This behaviour lowers formants beyond what would be expected based on mammalian acoustic allometry, thus exaggerating the animal's apparent size. Several researchers have proposed similar capabilities in humans, suggesting that systematic voice frequency modulation for size exaggeration should be observed not only across mammalian species³, but also across human cultures^{7,8}. Others have further hypothesized that exaggeration of body size through voice frequency modulation may have contributed to the descent of the human larynx⁹, and is likely to have played a critical role in the early evolution of nonverbal communication, ultimately paving the way for the emergence of articulated speech^{4,9}. The present study is the first to empirically test whether men or women do in fact systematically modulate F_0 and formants when instructed to deliberately alter their apparent body size.

Anatomical constraints on voice frequencies

Guided by the source-filter theory of speech production^{10,11}, behavioral scientists studying acoustic communication of body size in humans and other mammals have focused on two voice features: fundamental frequency (F_0) and vocal tract resonances (formants). Voice F_0 is produced by the vocal folds, whose rate of vibration is related to their mass, length and tension, whereas the supralaryngeal vocal tract filters the voice producing formants that are inversely related to supralaryngeal vocal tract length¹². Voice F_0 and formants affect our perception of pitch and timbre, respectively, and play a major role in speech articulation¹³. These voice features are also highly sexually dimorphic and have likely undergone intense sexual selection in humans¹⁴.

¹Department of Psychology, Neuroscience & Behaviour, McMaster University, Canada. ²Institute of Psychology, University of Wrocław, Poland. ³Mammal Vocal Communication & Cognition Research Group, School of Psychology, University of Sussex, United Kingdom. ⁴Department of Animal and Human Biology, Faculty of Biology, University of Havana, Cuba. ⁵Instituto de Ciencias Biomedicas, Universidad Autonoma de Chile, El Llano Subercaseaux 2801, San Miguel, Santiago, Chile. Correspondence and requests for materials should be addressed to K.P. (email: K.Pisanski@sussex.ac.uk)

Formants scale fairly allometrically with vocal tract length and body size¹⁵, because the mammalian vocal tract is constrained by the skeletal structures that surround it. In contrast, although larger vocal folds produce a lower F_0 , the larynx grows largely independently of the rest of the body and F_0 does not therefore scale allometrically with body size in humans¹⁶. Indeed, formants explain several times more variation in body size than does F_0 when sex and age are controlled¹⁷. Nevertheless, among humans, neither vocal feature explains a substantial portion of the variance in body size at the within-sex level^{17,18}.

The lack of a robust physical relationship between the human voice and body size suggests a lack of constraints to maintain allometry. Volitional voice modulation to exaggerate body size should therefore be possible, and could help to further explain this puzzling disassociation. At the perceptual level, and despite the lack of robust physical relationships, listeners cross-culturally associate both low F_0 and low formants with large body size even within sexes^{19–23}. This further suggests that similar to other mammals (see e.g. refs 24 and 25) the human voice conveys both honest and exaggerated cues to size. Perceptual correspondences between low voice frequencies and large body size are important because they may drive selection for vocal communication (or exaggeration) of size, even in the absence of robust physical relationships between the voice and body.

Morphological modifications for size exaggeration

The vocal anatomy of many mammals has undergone morphological modifications that appear to function, at least in part, to exaggerate apparent size¹. These include non-laryngeal velar vocal folds in koalas (*Phascolarctos cinereus*) that allow males to produce F_0 's typical of an animal as large as an elephant²⁶, the subhyoid air sacs in black-and-white colubus monkeys (*Colobus guereza*) that amplify resonant frequencies²⁴, and the descended larynx in males of several polygynous deer species⁶, and koalas²⁷, that enable them to produce low formant frequencies characteristic of much larger species.

Humans also have a descended larynx. In humans the descended larynx allows for the production of a broader range of speech sounds relative to the vocal repertoires of other primates²⁸, but importantly, also results in a lengthened pharyngeal cavity and thus relatively lower formants⁹. Among men, pubertal hormones cause the larynx to descend even further, a full vertebra lower than among women¹⁶, and cause men's vocal folds to grow 60% larger than women's²⁹. These morphological modifications are evolutionarily relevant, as they implicate a role of sexual selection and size exaggeration in the evolution of human vocal frequencies. However, men's F_0 and formants are approximately 80% and 20% lower than women's, respectively, and these sex differences in F_0 and formants exceed that which can be explained by sexual dimorphism in the vocal anatomy (i.e., men's vocal folds are on average only 60% larger than women's, and their vocal tracts are typically 15% longer) or by sexual dimorphism in body size (men are on average only 10% taller than women)³⁰. This discrepancy alludes to possible behavioural differences between men and women in vocal production or modulation³¹, wherein men may lower their F_0 and formants more than women through the behavioural mechanism of voice modulation. If true, voice modulation may account for some portion of the unexplained variance between men and women's vocal frequencies.

Voice frequency modulation in humans

Mechanistically, volitional modulation of F_0 is achieved by manipulating the tension and effective length or surface area of the vocal folds using the laryngeal muscles (cricothyroid muscles lengthen the vocal folds and increase F_0 , whereas thyroarytenoid muscles shorten the vocal folds and decrease F_0 , and their opposing effects can be coordinated or independent)^{32,33} or by increasing subglottal pressure. In contrast, lowering the larynx or protruding the lips increases supralaryngeal vocal tract length and reduces formant spacing^{13,32,33}. Although recent investigations suggest some flexible control of voice frequencies in nonhuman primates^{34–36}, the ability to intentionally and volitionally modulate source and filter components is uniquely advanced in humans and is thought to constitute a precursor of speech^{4,9}. Indeed volitional voice modulation in humans involves comparatively complex neural processes that are absent in other mammals, including nonhuman primates³⁷.

Infant directed speech, in which adults speak with higher F_0 and exaggerated prosodic cues when addressing infants compared to older individuals, represents perhaps the most extensively studied form of voice modulation in humans and appears to be present across diverse cultures^{38,39}. More recently, a small number of empirical studies have begun to examine voice modulation as a social tool used to exploit ecologically relevant traits, and among these, almost all have focused on F_0 modulation (see ref. 4 for review). For example, in a series of recent studies, Cartei and colleagues^{40–42} showed that men, women, and children volitionally decreased both F_0 and formants when asked to sound masculine, and increased both voice features to sound feminine. Several studies report F_0 modulation in men or women when speaking to a potential mate^{43–48} or competitor⁴⁹. In the context of mate preferences, these studies have found that both sexes volitionally modulate F_0 when instructed to speak in a more attractive voice⁴³ as well as when directing their speech toward an attractive person of the opposite sex^{45,47}.

Voice modulation may therefore be utilized to deemphasize or accentuate various indexical traits and this may be evolutionary adaptive. In particular, men who can effectively exaggerate their apparent body size through F_0 and formant modulation may reap the social benefits associated with physical largeness, such as increased access to resources and mates. Indeed, taller men, and those with relatively lower voice F_0 and formants indicating larger body size, are typically preferred as mates by women across a diverse range of cultures⁵⁰. Nevertheless, to be effective, vocal modulation of body size should exceed the just-noticeable differences in F_0 /formant perception^{23,51,52} and should have the intended effects on listeners' social assessments. While some studies have found that volitional voice modulation effectively increased listeners' assessments of the vocalizer's attractiveness, competence, and intelligence^{43,47}, one study found that sex-typical F_0 modulation influenced listeners' assessments of dominance but not voice attractiveness⁴⁶.

		Formant Modulation		Voice Pitch Modulation			
		VTL, cm (%)		F0, Hz (%)		F0, ERB (%)	
		Men	Women	Men	Women	Men	Women
Large Condition	Mean	0.92 (5.3%)	0.54 (3.6%)	-1.03 (0.8%)	-8.57 (4%)	-0.03 (0.8%)	-0.19 (3.1%)
	Maximum	4.00 (23.1%)	3.01 (20.3%)	-43.36 (35.5%)	-75.00 (35.1%)	-1.05 (26.5%)	-1.69 (27.6%)
Small Condition	Mean	-0.51 (3.0%)	-0.20 (1.4%)	32.06 (26.3)	24.80 (11.6%)	0.73 (18.4)	0.48 (7.8%)
	Maximum	-2.79 (16.1%)	-1.65 (11.1%)	345.17 (282.6%)	138.64 (64.8%)	6.30 (159.0%)	2.70 (44.1%)

Table 1. Means and maxima in VTL (cm) and F0 modulation (Hz and ERB) for each sex and condition, given in absolute units and percentage change from baseline.

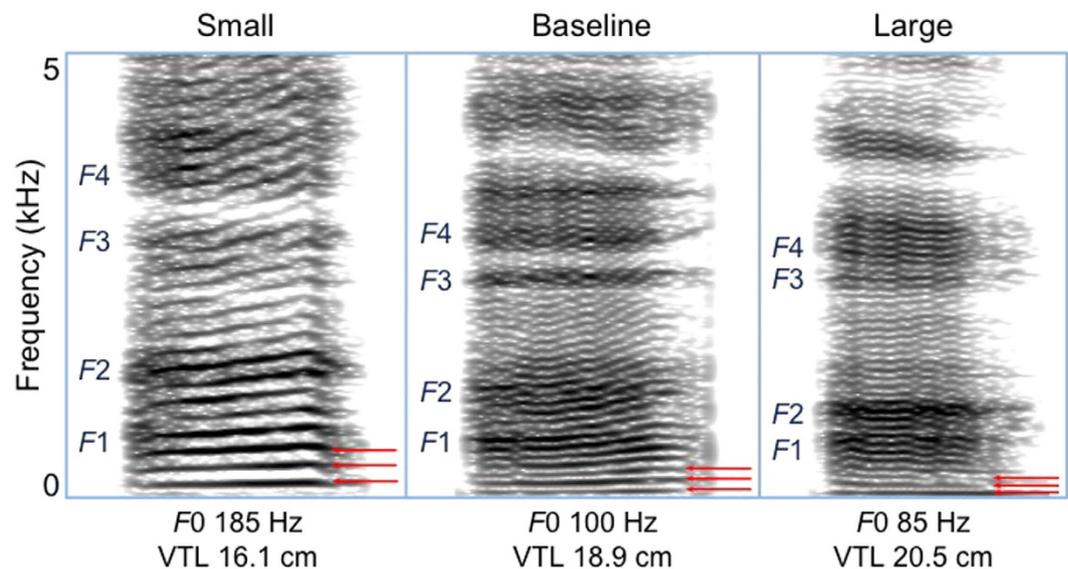


Figure 1. Spectrograms illustrating the vowel /a/ spoken by the same adult male in each condition.

Formants (F1–F4) are labeled. Fundamental frequency and the first two harmonics (multiple integers of F0) are indicated by red arrows. Participants raised formants and F0 to sound small, thus increasing spacing between F1–F4 and between harmonics (left), and lowered formants and F0 to sound large (right). Gaussian FFT, window length 0.04; dynamic range 60 dB. Refer to Supplementary Audio S1 for corresponding voice recording.

The Present Study

The present study is the first to test whether humans can modulate voice features known to be associated with body size (fundamental and formant frequencies) when instructed to deliberately alter their apparent body size. In addition, we examined whether this voice modulation reflects real (physical) and perceived relationships between the human voice and body (i.e., lower F0 and formants indicate larger size and *visa versa*), whether the behaviour differs between the sexes, and whether the behaviour is present cross-culturally.

We tested these hypotheses in 167 men and women from three distinct cultures and language groups: Canada (English), Cuba (Spanish), and Poland (Polish). Participants were recorded speaking vowel sounds in their baseline voice and while imitating a physically large and small body size. We predicted that participants would lower F0 and formants (increase apparent vocal tract length, VTL) to convey large size, and raise voice F0 and formants (reduce VTL) to convey small size. We further predicted that men would modulate their voices more than women, thereby accounting for some of the unexplained sexual dimorphism in F0 and formants. In contrast, we predicted that patterns of voice modulation would not differ across the three cultures. This latter finding would provide some support for fairly universal sound-size correspondences, and/or anatomical or biomechanical constraints on voice modulation.

The present study was specifically designed to test for the first time whether adult speakers are capable of volitional adjustments to their larynx (fundamental frequency modulation) and vocal tract (formant frequency modulation) in a manner that parallels the known relationships between these vocal parameters and body size in humans. Acoustic analyses were utilized to measure voice frequency parameters and to test whether these modulations exceed just-noticeable differences in F0 and formant perception. However in the present study we did not test whether these modulations effectively alter listeners' perceptions of the vocalizer's body size.

Results

Table 1 shows unstandardized means and maxima in VTL and F0 modulation for each sex and condition. As predicted, both sexes decreased VTL and increased F0 to sound small, and increased VTL and decreased F0 to

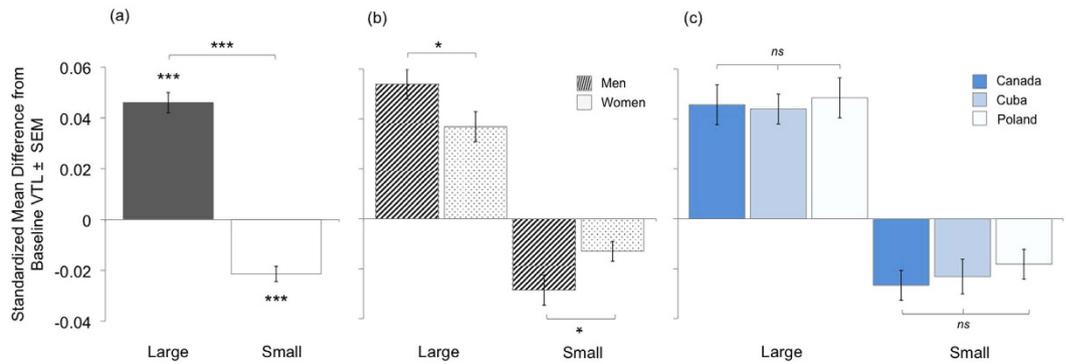


Figure 2. Vocal tract length (VTL) modulation given as the standardized difference from baseline in the large and small conditions. (a) Participants increased VTL to sound physically large and decreased VTL to sound small. The magnitude of VTL modulations was greater in the large than small condition. (b) Men modulated their VTLs more than did women in both size conditions. (c) VTL modulation did not vary cross-culturally. *** $p < 0.001$, * $p < 0.05$, ns $p > 0.05$.

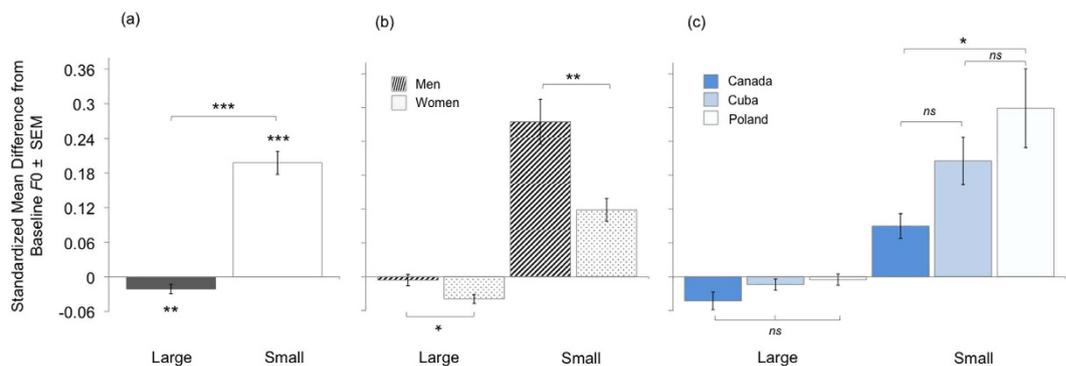


Figure 3. Fundamental frequency (F0) modulation. (a) Participants decreased F0 to sound physically large, and increased F0 to sound small. The magnitude of F0 modulations was greater in the small than large condition. (b) Men modulated VTL more than did women, but only in the small condition. (c) F0 modulation did not vary cross-culturally in the large condition, however in the small condition, Poles modulated their F0 more than did Canadians. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$.

sound large (Fig. 1; Supplementary Audio S1). Notably, men increased their apparent VTLs by as much as 25% to portray a physically larger body size, and increased their F0 by up to three times the baseline frequency (i.e., almost 300%) to sound smaller, reaching pitch registers characteristic of a child⁵³.

Formant or vocal tract length modulation. An analysis of variance revealed a main effect of condition (large versus small body size imitation) ($F_{1,111} = 109.2$, $p < 0.001$, $\eta_p^2 = 0.50$; Fig. 2a) and an interaction between condition and sex ($F_{1,111} = 8.1$, $p = 0.005$, $\eta_p^2 = 0.07$; Fig. 2b) on VTL modulation. There were no other significant effects (all $F < 2.1$, all $p > 0.13$) including no effects of culture (Fig. 2c). Post-hoc analyses showed that participants increased their VTL from baseline in the large condition (one-sample $t_{132} = 9.7$, $p < 0.001$) and decreased their VTL in the small condition ($t_{132} = -5.4$, $p < 0.001$). Moreover, men increased VTL in the large condition (one-way $F_{1,132} = 6.01$, $p = 0.016$) and decreased VTL in the small condition ($F_{1,122} = 5.78$, $p = 0.018$) significantly more than did women. A model examining *absolute* differences from baseline (i.e., magnitude of modulation) indicated that VTL modulations were more extreme in the large than small condition, and more extreme among men than women in both conditions (see Supplementary Information; see also Fig. 2).

Fundamental frequency modulation. We observed main effects of condition ($F_{1,161} = 55.77$, $p < 0.001$, $\eta_p^2 = 0.26$; Fig 3a), sex ($F_{1,161} = 10.7$, $p = 0.001$, $\eta_p^2 = 0.06$; Fig 3b) and culture ($F_{2,161} = 6.1$, $p = 0.003$, $\eta_p^2 = 0.07$; Fig 3c) on F0 modulation. These effects were qualified by a significant interaction between condition and sex ($F_{2,161} = 4.4$, $p = 0.037$, $\eta_p^2 = 0.03$) and a marginally non significant interaction between condition and culture ($F_{2,161} = 3.1$, $p = 0.051$, $\eta_p^2 = 0.04$). There were no other significant effects (all $F < 1.9$, all $p > 0.16$).

Planned post-hoc analyses showed that participants decreased their F0 in the large condition (one-sample $t_{166} = -2.6$, $p = 0.01$) and increased their F0 in the small condition ($t_{166} = 6.7$, $p < 0.001$). Men increased their F0 more than did women to sound small (one-way $F_{1,166} = 7.2$, $p = 0.008$), however women decreased their F0 more than did men to sound large ($F_{1,166} = 5.5$, $p = 0.021$). Cultural differences in F0 modulation emerged only in the

small condition ($F_{2,166} = 4.4, p = 0.014$), and only between Canadians and Poles (Fisher's LSD $p = 0.004$; all other $p > 0.11$; Fig. 3c). A model examining absolute magnitude indicated that $F0$ modulations were more extreme in the large than small condition. Within the small condition, $F0$ modulations were more extreme among men than women (see Supplementary Information; see also Fig. 3).

Discussion

The capacity for humans to volitionally modulate the source and filter components of our voices has traditionally been studied in the context of speech and language production^{9,11}. The extent to which we modulate our voices for nonverbal communication, for instance to sound more masculine/feminine or attractive, has been investigated in comparatively few empirical studies^{40,41,43–48,54,55}. Our study provides the first evidence that men and women from diverse cultures can spontaneously and volitionally modulate their fundamental and formant frequencies with the intent to exaggerate or reduce apparent body size, and that regardless of culture, men generally modulate their voices more than do women in this context. Acoustic analyses indicated that these modulations were in the predicted direction, such that men and women lowered $F0$ and formants when instructed to sound large, and increased $F0$ and formants when instructed to sound small, and that in most cases these modulations exceeded the just-noticeable differences in $F0$ and formant perception.

The patterns of voice frequency modulation observed in our study map onto real physical relationships between the voice and body, as larger people generally have lower formants and $F0$ than do smaller people^{17,18,22}. However, because neither vocal parameter (especially $F0$) can explain a substantial proportion of the variance in human body size when sex and age are controlled^{17,18,22}, volitional voice modulation of these parameters may also reflect an exploitation of listeners' perceptual biases linking low voice frequencies to large body size and dominance^{7,8,21–23,54}, or more general sound symbolic correspondences⁵⁶. Indeed our results support Ohala's prediction that similar voice frequency modulations will be observed across cultures, reflecting a universal "frequency code"^{9,7,8}. It has also previously been suggested that perceptual biases based on the laws of physics, such that large objects resonate at lower frequencies, are likely to be cross-culturally universal precisely because they are determined by physics, not culture⁵⁷ (see also ref. 3). Our cross-cultural results may alternatively reflect constraints on voice production in humans. Formants are especially constrained by the bony anatomy surrounding the vocal tract¹⁵, which is likely to impose upper and lower limits on formant modulation.

The sex differences in voice modulation observed here may be tied to a number of factors, most parsimoniously to differences in the vocal anatomy of men and women. For example, a longer supralaryngeal vocal tract among men may allow for greater laryngeal mobility that could result in a broader range of formant manipulations. Men's voices are also lower in frequency than are women's, and as a result men must raise their voices more than women to reach similar high frequency targets. Nevertheless our results indicate that men exceeded the frequency targets reached by women even when raising their voice frequencies to sound small. Indeed we observed extreme maxima in modulations of both $F0$ and VTL, particularly among men. On one hand, this demonstrates an impressive capacity for men to volitionally manipulate their larynges and vocal tracts. On the other hand, it elicits a question about the ecological validity of such extreme modulations, which may be perceived as abnormal.

Our results indicate that speakers modulated $F0$ more than VTL. We also observed asymmetries within each vocal parameter, specifically greater decreases than increases in formants, and greater increases than decreases in $F0$. This latter finding might be explained by nonlinearities in the relationship between vocal fold length and $F0$ ³², and the greater physiological effort required to increase versus decrease vocal fold tension⁵⁸. Indeed baseline $F0$ is closer to the minimum than maximum producible $F0$ ¹². As a consequence, sopranos can reach $F0$'s above 1200 Hz, whereas bass singers lower their $F0$ by only a fraction of this magnitude, typically to around 80 Hz⁵⁹.

The demonstrated capacity to volitionally modulate vocal parameters known to be physically related to and perceptually associated with body size can be evolutionarily advantageous, as various indicators of physical size in humans are known to influence a wide range of socioeconomic variables and the mate preferences of both sexes⁵⁰. At the same time, voice modulation is ecologically relevant only if and when it affects listeners. Perceptually, human listeners can discriminate changes in $F0$ or formants of about 5% from a series of vowel sounds⁵², and formant manipulations of 5% are known to affect listeners' body size estimates⁶⁰. Based on this our results suggest that, on average, men's formant-based size exaggeration, and both men's and women's $F0$ -based size reduction, would be perceptually detectable. Studies examining the effectiveness of voice modulation on other types of judgments have produced mixed results^{43,46,47}, but generally suggest that voice modulation may be an effective tool for manipulating listeners' social judgments of traits such as attractiveness, dominance, and competence. For instance, one recent study found that listeners preferred the voices of men and women whose speech was directed towards attractive individuals, and these preferences were observed for voices recorded in the listener's own language as well as in a foreign language⁴⁷. In the case of vocally faking a larger body size, and thus a more dominant persona, individuals who are perceived as physically larger due to voice modulation could reap the socioeconomic and reproductive benefits typically linked to these traits across various social contexts including mating, political and marketing contexts. Currently we are conducting playback experiments to test whether vocal modulation can effectively alter listeners' estimates of body size.

Methods

Participants. A total of 167 men and women from Canada (students of McMaster University in Hamilton), Cuba (students of the University of Havana, and staff and students of the Cuban Neuroscience Center in Havana), and Poland (students of the University of Wrocław and the College of Humanities and Economics in Brzeg) took part in the experiment. All participants provided informed consent. Sample characteristics are given in Table 2.

Procedure. All participants were first recorded speaking the five monophthong vowels /a/, /i/, /ε/, /o/, and /u/ (International Phonetic Alphabet) in their natural, baseline voice. Following this, participants were asked to

Country	Sex	<i>n</i>	Age, years	Height, cm	Weight, kg	<i>F</i> ₀ , Hz	VTL, cm
Canada	Men	28	19.7 (2.3, 18–30)	177.4 (7.1)	73.4 (11.3)	114.1 (14.6)	17.3 (0.9)
	Women	28	20 (1.4, 18–25)	164.6 (6.8)	60.0 (12.1)	218.2 (22.2)	14.7 (0.6)
Cuba	Men	30	25.9 (2.9, 21–32)	174.6 (6.9)	73.2 (9.6)	131.8 (21.0)	17.4 (0.7)
	Women	25	25.8 (3.2, 21–32)	160.2 (4.6)	58.2 (11.4)	211.7 (18.3)	14.7 (0.6)
Poland	Men	27	28.4 (5.5, 22–42)	181.1 (6.1)	81.3 (11.5)	120.0 (14.7)	17.4 (0.9)
	Women	29	26.4 (6.5, 21–45)	165.1 (5.9)	67.8 (17.1)	211.9 (23.2)	15.2 (0.6)

Table 2. Sample characteristics (mean (*s.d.*, range)).

repeat the five vowels while sounding physically small (small condition) and physically large (large condition). These instructions, back translated and given in the native language of the participant, were the only instructions given. Condition order was counter-balanced between participants. Participants then completed a short questionnaire indicating their sex and age. Height was measured using metric tape and weight using an electronic scale. The study was approved by the McMaster Research Ethics Board and methods were carried out in accordance with the approved guidelines.

Voice recording. All participants were recorded using condenser microphones with a cardioid pick-up pattern at an approximate distance of 5–10 cm (Canada: Sennheiser MKH 800; Cuba: Sennheiser MKH 70; Poland: Audio-M Nova). Audio was digitally encoded with an M-Audio Fast Track interface at a sampling rate of 44.1–96 kHz and 16–24 bit amplitude quantization, and stored onto a computer as PCM WAV files. Recordings from participants at McMaster University and the Cuban Neuroscience Center were conducted in an anechoic sound-controlled booth and recordings at the Universities of Havana and Wrocław were conducted in a quiet room.

Voice measurement and acoustic analysis. All acoustic measures were performed in Praat⁶¹. Voice measures were taken from each vowel separately and then averaged across vowels within each vocalizer and condition to obtain mean values. We measured *F*₀ using Praat's autocorrelation algorithm. Following previous work, we set a broad search range of 30–500 Hz for men, and 65–600 Hz for women⁴¹. We transformed *F*₀ measures into equivalent rectangular bandwidth (ERB) units, a quasi-logarithmic scale that controls for the difference between physical and perceived properties of pitch, where 1 ERB is approximately equal to a 40 Hz change at a centre frequency of 120 Hz⁶². The ERB scale correlates strongly with *F*₀ in Hz in the range of adult human speech (e.g., $r = 0.99$ in men)²¹.

We measured formants (*F*₁–*F*₄) using Praat's Burg Linear Predictive Coding algorithm with the initial settings of maximum formant set to 5500 Hz for women and 5000 Hz for men. Formants were first overlaid on a spectrogram and formant number was manually adjusted until the best visual fit of predicted onto observed formants was obtained. From the mean centre frequencies of *F*₁–*F*₄ we computed formant spacing, ΔF , a measure of the distance among adjacent formants, as well as apparent vocal tract length derived from formant spacing, $VTL(\Delta F)$ ⁶³. The results of a recent meta-analysis indicate that ΔF and $VTL(\Delta F)$ each independently explain more variance in men's heights and women's weights than do any other formant measures¹⁷, and are strongly inversely related (here, $r = -0.99$ within each sex).

Each individual formant is related to ΔF by Equation (1):

$$F_i = \frac{(2i - 1)}{2} \Delta F \quad (1)$$

where *i* represents formant position (*F*₁–*F*₄). Thus, we derived ΔF by plotting mean formant frequencies for each individual against the expected increments of formant spacing $[(2i - 1)/2]$, where ΔF is equal to the slope of the linear regression line with an intercept set to 0^{41,63}. From this, we estimated the apparent vocal tract length of each individual following equation (2):

$$VTL(\Delta F) = \frac{c}{2\Delta F} \quad (2)$$

where *c* is 35 000 cm/s, the approximate speed of sound in a uniform tube with one end closed controlling for warmth and dampness (i.e. the vocal tract¹²). From the pooled samples, we confirmed that baseline VTL explained several times more variance in men's (12%, $r_s = 0.35$) and women's (16%, $r_s = 0.40$) heights than did baseline *F*₀ (2.5% in each sex, $r_s = 0.16$; See Supplementary Fig. S1). This pattern of results was similar across samples and agrees with weighted relationships reported at the population level¹⁷.

Statistical analysis. We first calculated differences in voice measures between each size condition and baseline, separately for *F*₀ and VTL. Positive values indicate increases, and negative values decreases, from baseline. We then ran separate repeated measures ANOVAs for *F*₀ and VTL. In each model, the dependent variable was the standardized difference from baseline ($[\text{large} - \text{baseline}] / \text{baseline}$; $[\text{small} - \text{baseline}] / \text{baseline}$), controlling for baseline sex differences. Condition (large, small) was included as a within-subject factor, and sex (male, female) and culture (Canada, Cuba, Poland) as between-subject factors. To examine differences in the *magnitude* of voice modulations, we re-ran the models on the absolute standardized difference from baseline in each condition

(see Supplementary Information). Significant effects were further examined using planned post-hoc tests. All tests were two-tailed with an alpha of 0.05.

References

- Fitch, W. T. & Hauser, M. D. In *Acoustic Communication* (eds. Simmons, A. M., Fay, R. R. & Popper, A. N.) 65–137 (Springer New York, 2003).
- Taylor, A. M. & Reby, D. The contribution of source-filter theory to mammal vocal communication research: Advances in vocal communication research. *J. Zool.* **280**, 221–236 (2010).
- Morton, E. S. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* **111**, 855–869 (1977).
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J. & Reby, D. Voice modulation: A window into the origins of human vocal control? *Trends Cogn. Sci.* (2016).
- Reby, D. *et al.* Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proc. R. Soc. Lond. B Biol. Sci.* **272**, 941–947 (2005).
- Fitch, W. T. & Reby, D. The descended larynx is not uniquely human. *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 1669–1675 (2001).
- Ohala, J. J. Cross-language use of pitch: an ethological view. *Phonetica* **40**, 1–18 (1983).
- Ohala, J. J. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* **41**, 1–16 (1984).
- Fitch, W. T. The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267 (2000).
- Chiba, T. & Kajiyama, M. *The vowel: Its nature and structure*. (Phonetic Society of Japan, 1958).
- Fant, G. *Acoustic theory of speech production*. (Mouton, 1960).
- Titze, I. R. *Principles of vocal production*. (Prentice-Hall, 1994).
- Kreiman, J. & Sidtis, D. *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. (Wiley-Blackwell, 2011).
- Puts, D., Jones, B. C. & DeBruine, L. M. Sexual selection on human faces and voices. *J. Sex Res.* **49**, 227–243 (2012).
- 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).
- Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M. & Palmer, J. B. Ontogeny of postnatal hyoid and larynx descent in humans. *Arch. Oral Biol.* **46**, 117–128 (2001).
- Pisanski, K. *et al.* Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* **95**, 89–99 (2014).
- Pisanski, K. *et al.* Voice parameters predict sex-specific body morphology in men and women. *Anim. Behav.* **112**, 13–22 (2016).
- 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).
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M. & Perrett, D. I. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim. Behav.* **69**, 561–568 (2005).
- Pisanski, K., Fraccaro, P. J., Tighe, C. C., O'Connor, J. J. M. & Feinberg, D. R. Return to Oz: Voice pitch facilitates assessments of men's body size. *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 1316–1331 (2014).
- Rendall, D., Vokey, J. R. & Nemeth, C. Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 1208–1219 (2007).
- Smith, D. R. R. & Patterson, R. D. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and agea. *J. Acoust. Soc. Am.* **118**, 3177–3186 (2005).
- Harris, T. R., Fitch, W. T., Goldstein, L. M. & Fashing, P. J. Black and White Colobus Monkey (*Colobus guereza*) Roars as a Source of Both Honest and Exaggerated Information About Body Mass. *Ethology* **112**, 911–920 (2006).
- Sanvito, S., Galimberti, F. & Miller, E. H. Vocal signalling of male southern elephant seals is honest but imprecise. *Anim. Behav.* **73**, 287–299 (2007).
- Charlton, B. D. *et al.* Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Curr. Biol.* **23**, R1035–R1036 (2013).
- 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).
- Lieberman, P. H., Klatt, D. H. & Wilson, W. H. Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science* **164**, 1185–1187 (1969).
- Kahane, J. C. A morphological study of the human prepubertal and pubertal larynx. *Am. J. Anat.* **151**, 11–19 (1978).
- Titze, I. R. Physiologic and acoustic differences between male and female voices. *J. Acoust. Soc. Am.* **85**, 1699–1707 (1989).
- Rendall, D., Kollias, S., Ney, C. & Lloyd, P. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *J. Acoust. Soc. Am.* **117**, 944 (2005).
- Titze, I. R. Vocal fold mass is not a useful quantity for describing f0 in vocalization. *J. Speech Lang. Hear. Res.* **54**, 520–522 (2011).
- Hollien, H. Vocal fold dynamics for frequency change. *J. Voice* **28**, 395–405 (2014).
- Clay, Z., Archbold, J. & Zuberbühler, K. Functional flexibility in wild bonobo vocal behaviour. *PeerJ* **3**, e1124 (2015).
- Hotchkiss, C. F., Parks, S. E. & Weiss, D. J. Noise-Induced Frequency Modifications of Tamarin Vocalizations: Implications for Noise Compensation in Nonhuman Primates. *PLoS ONE* **10**, e0130211 (2015).
- Lameira, A. R. *et al.* Speech-like rhythm in a voiced and voiceless orangutan call. *Plos One* **10**, e116136 (2015).
- 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).
- Falk, D. Prelinguistic evolution in early hominins: Whence motherese? *Behav. Brain Sci.* **27**, 491–541 (2004).
- Bryant, G. A. & Barrett, H. C. Recognizing intentions in infant-directed speech: Evidence for universals. *Psychol. Sci.* **18**, 746–751 (2007).
- Cartei, V. & Reby, D. Acting gay: Male actors shift the frequency components of their voices towards female values when playing homosexual characters. *J. Nonverbal Behav.* **36**, 79–93 (2011).
- Cartei, V., Cowles, H. W. & Reby, D. Spontaneous voice gender imitation abilities in adult speakers. *Plos One* **7**, e31353 (2012).
- Cartei, V. & Reby, D. Effect of formant frequency spacing on perceived gender in pre-pubertal children's voices. *Plos One* **8**, e81022 (2013).
- Hughes, S. M., Mogilski, J. K. & Harrison, M. A. The perception and parameters of intentional voice manipulation. *J. Nonverbal Behav.* **38**, 107–127 (2014).
- Hughes, S. M., Farley, S. D. & Rhodes, B. C. Vocal and physiological changes in response to the physical attractiveness of conversational partners. *J. Nonverbal Behav.* **34**, 155–167 (2010).
- Fraccaro, P. J. *et al.* Experimental evidence that women speak in a higher voice pitch to men they find attractive. *J. Evol. Psychol.* **9**, 57–67 (2011).
- Fraccaro, P. J. *et al.* Faking it: deliberately altered voice pitch and vocal attractiveness. *Anim. Behav.* **85**, 127–136 (2013).
- Leongómez, J. D. *et al.* Vocal modulation during courtship increases perceptivity even in naive listeners. *Evol. Hum. Behav.* **35**, 489–496 (2014).

48. Anolli, L. & Ciceri, R. Analysis of the vocal profiles of male seduction: from exhibition to self-disclosure. *J. Gen. Psychol.* **129**, 149–169 (2002).
49. Puts, D., Gaulin, S. J. C. & Verdolini, K. Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* **27**, 283–296 (2006).
50. Pisanski, K. & Feinberg, D. R. Cross-cultural variation in mate preferences for averageness, symmetry, body size, and masculinity. *Cross-Cult. Res.* **47**, 162–197 (2013).
51. Re, D. E., O'Connor, J. J. M., Bennett, P. J. & Feinberg, D. R. Preferences for very low and very high voice pitch in humans. *PLoS ONE* **7**, e32719 (2012).
52. 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).
53. Baken, R. J. & Orlikoff, R. F. *Clinical Measurement of Speech and Voice*. (Cengage Learning, 2000).
54. Puts, D., Hodges, C. R., Cárdenas, R. A. & Gaulin, S. J. C. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* **28**, 340–344 (2007).
55. Cartei, V., Cowles, W., Banerjee, R. & Reby, D. Control of voice gender in pre-pubertal children. *Br. J. Dev. Psychol.* **32**, 100–106 (2014).
56. Hinton, L., Nichols, J. & Ohala, J. J. *Sound Symbolism*. (Cambridge University Press, 2006).
57. Spence, C. Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* **73**, 971–995 (2011).
58. Traunmüller, H. & Eriksson, A. The frequency range of the voice fundamental in the speech of male and female adults. *Unpubl. Manuscr.* (1995).
59. Joliveau, E., Smith, J. & Wolfe, J. Acoustics: Tuning of vocal tract resonance by sopranos. *Nature* **427**, 116–116 (2004).
60. Irino, T., Aoki, Y., Kawahara, H. & Patterson, R. D. Comparison of performance with voiced and whispered speech in word recognition and mean-formant-frequency discrimination. *Speech Commun.* **54**, 998–1013 (2012).
61. Boersma, P. & Weenink, D. *Praat: Doing phonetics by computer*. (2015).
62. Traunmüller, H. Analytical expressions for the tonotopic sensory scale. *J. Acoust. Soc. Am.* **88**, 97–100 (1990).
63. 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).

Acknowledgements

The authors thank Maydel Fernandez, Lida Sánchez, Nadir Díaz Simón, and Joanna Widomska for assisting in data collection. This research was made possible by a Michael Smith Foreign Study Supplement in Cuba from the Social Sciences and Humanities Research Council of Canada (771-2013-0108), as well as by funding from the National Science Center (2014/13/B/HS6/02636), the Foundation for Polish Science, the University of Wrocław, and a Marie-Skłodowska Curie Individual Fellowship (H2020-16 MSCA-IF-2014-655859) to KP.

Author Contributions

Authors K.P., D.R. and D.R.F. made substantial theoretical contributions. All authors contributed conceptually to the experimental design and collected cross-cultural data; K.P. performed acoustic measures, statistical analyses and drafted the manuscript and figures. All authors revised the paper for content and approved the final version for submission.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Pisanski, K. *et al.* Volitional exaggeration of body size through fundamental and formant frequency modulation in humans. *Sci. Rep.* **6**, 34389; doi: 10.1038/srep34389 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016