Voice of Authority: Professionals Lower Their Vocal Frequencies When Giving Expert Advice

Piotr Sorokowski¹ · David Puts^{2,3,4} · Janie Johnson² · Olga Żółkiewicz¹ · Anna Oleszkiewicz^{1,5} · Agnieszka Sorokowska^{1,6} · Marta Kowal¹ · Barbara Borkowska⁷ · Katarzyna Pisanski^{1,8}

Published online: 7 May 2019
© The Author(s) 2019

Abstract

Acoustic analysis and playback studies have greatly advanced our understanding of between-individual differences in nonverbal communication. Yet, researchers have only recently begun to investigate within-individual variation in the voice, particularly how people modulate key vocal parameters across various social contexts, with most of this research focusing on mating contexts. Here, we investigated whether men and women modulate the frequency components of their voices in a professional context, and how this voice modulation affects listeners' assessments of the speakers' competence and authority. Research assistants engaged scientists working as faculty members at various universities in two types of speech conditions: (1) Control speech, wherein the subjects were asked how to get to the administrative offices on that given campus; and (2) Authority speech, wherein the same subjects were asked to provide commentary for a radio program for young scholars titled, "How to become a scientist, and is it worth it?". Our results show that male ($n=27$) and female ($n=24$) faculty members lowered their mean voice pitch (measured as fundamental frequency, F_0) and vocal tract resonances (measured as formant position, F_p) when asked to provide their expert opinion compared to when giving directions. Notably, women lowered their mean voice pitch more than did men (by 33 Hz vs. 14 Hz) when giving expert advice. The results of a playback experiment further indicated that foreign-speaking listeners judged the voices of faculty members as relatively more competent and more authoritative based on authority speech than control speech, indicating that the observed nonverbal voice modulation effectively altered listeners' perceptions. Our results support the prediction that people modulate their voices in social contexts in ways that are likely to elicit favorable social appraisals.

Keywords Authority · Fundamental frequency · Voice pitch · Formant frequencies · Voice modulation

✉ Katarzyna Pisanski
kasiapisanski@gmail.com

Extended author information available on the last page of the article

Introduction

Nonhuman animals have long been known to communicate dominance and aggression through their vocalizations (Morton 1977). Lower frequency vocalizations are typically produced by relatively larger, more dominant individuals and are used by many animals to signal threat (Taylor and Reby 2010). Humans are no exception. Men's fundamental frequency (F_0 , perceived as voice pitch) and vocal tract resonances or formant frequencies are significantly lower than women's due mostly to sexual dimorphism in human vocal anatomy (Titze 1989). However, even at the within-sex level, larger or more dominant men and women typically produce lower fundamental and formant frequencies (i.e., lower pitched and more resonant speech) relative to their smaller, subordinate counterparts (Puts et al. 2012a). Listeners are in turn highly sensitive to these correspondences. In fact, in addition to associating low-frequency voices with greater physical size, strength, masculinity, and dominance (reviewed in Pisanski and Bryant 2019), listeners often associate low voice pitch with positive psychological traits such as competence and intelligence (Kreiman and Sidtis 2011 for review; cf. Hughes et al. 2014). These vocal stereotypes can have real-life implications affecting, for instance, who people prefer to vote for in a political election (Gregory Jr. and Gallagher 2002; Klofstad et al. 2015; Tigue et al. 2012) or to hire following a job interview (Schroeder and Epley 2015).

Men with relatively lower voice pitch and perhaps lower formants are often also judged as more attractive compared to men with relatively higher voice frequencies (Pisanski and Feinberg 2019 for review; Puts et al. 2016). In a professional context, both men and women with relatively low voice frequencies are typically judged as more dominant and competent (see e.g., Klofstad et al. 2012). Thus, a low-frequency voice may benefit men in a broad range of social contexts ranging from sexual to political and economic. This is not always the case among women, for whom, like men, low voice frequencies are perceived as masculine (Pisanski and Feinberg 2019), but can be considered unattractive (Feinberg 2008; Puts 2016). Indeed, most studies that have either experimentally manipulated voice frequencies or examined natural variation in nonverbal vocal parameters found that deeper voices are perceived as less attractive in women (Pisanski and Feinberg 2019; Puts et al. 2016). Nevertheless, several studies have found the opposite (Babel et al. 2014; Hughes et al. 2014; Pisanski et al. 2018; Tuomi and Fisher 1979), perhaps because deeper voices may signal sexual interest, maturity, confidence, and/or competence.

Humans can readily and volitionally modulate various nonverbal vocal parameters including their fundamental and formant frequencies (Pisanski et al. 2016a). Yet, despite a growing number of studies highlighting the social relevance of between-individual differences in voice production, few studies have examined vocal variation within individuals. The capacity for voice modulation in humans, in conjunction with evidence for strong vocal stereotypes that can have real life implications, poses the question of whether men and women modulate the frequency components of their voices in specific social contexts to elicit a desired social appraisal or outcome. Indeed, the rare capacity for humans (compared to other animals) to modulate the nonverbal parameters of the voice may have been selected for, and may confer various evolutionary and social advantages (Pisanski et al. 2016a).

A small number of lab-based studies examining voice modulation in socially relevant contexts, most within in a mating context, provide preliminary support for this hypothesis. For example, Puts et al. (2006) demonstrated voice pitch modulation among men taking part in a laboratory-based mock dating game. Men who rated

a male competitor as more dominant relative to themselves raised their voice pitch compared to a control condition, whereas men who rated the competitor as less dominant lowered their voice pitch (± 2 Hz on average). Using the same dataset, Hodges-Simeon et al. (2010) found that the standard deviation of men's fundamental frequency changed from baseline when speaking to a potential date as a function of the man's self-rated relative dominance. Subsequent research has revealed that both sexes generally modulate their voice pitch when speaking to attractive members of the opposite sex (Fraccaro et al. 2011; Hughes et al. 2010, 2014; Leongómez et al. 2014). Pisanski et al. (2018) recently showed that men and women change their voice pitch during real life speed dating, depending both on whether they show a personal preference for their date, and on his or her overall desirability. Finally, Leongómez et al. (2017) demonstrated that interviewees modulated the pitch of their voices during mock interviews, raising their voice pitch in response to employers they perceived as dominant and prestigious. Several laboratory studies have also demonstrated that men and women can volitionally modulate their voices "on demand", for example, when instructed to sound 'confident, intelligent or dominant' (Hughes et al. 2014), 'masculine or feminine' (Cartei et al. 2012), or 'physically large or small' (Pisanski et al. 2016c).

Yet, there remains limited evidence for context-dependent voice modulation in real-life social contexts, outside of the laboratory, where the mechanisms and outcomes may differ compared to voice modulation observed in more controlled settings. Much research has related voice modulation to the emotional content of speech (Scherer 2019 for recent review). Other research has sought to identify the patterns and natural sequences of voice modulations occurring when people interact. For instance, nonverbal vocal parameters can change within the same individual across the span of a single conversation or interview (Gregory et al. 1993), during oral examinations (Pisanski et al. 2016d), when lying (Ekman et al. 1976), or depending on the conversational partner (e.g., infant-directed speech: Wang et al. 2019). It is not entirely clear how much of these vocal changes are a consequence of emotional arousal, or a result of conscious voice modulation.

Anecdotal evidence also suggests that people manipulate their voices when taking a position of authority. For example, former British Prime Minister Margaret Thatcher reportedly lowered her voice pitch when delivering political speeches (Karpf 2006). However, very few studies have directly tested this hypothesis. Hunter and Titze (2010) examined teachers' vocal intensity (voicing percentage per hour, loudness – dB SPL and fundamental frequency) during and after work hours. Occupational voicing percentage per hour was more than twice that of non-occupational, and occupational voices were higher in pitch. However, several elements of the study design (e.g., no experimental conditions, lack of control group, increased noise in a classroom as compared to home environment) limit firm conclusions regarding voice modulation when speaking authoritatively.

The present study was therefore designed to test (a) whether professional men and women alter their voice frequencies when speaking authoritatively about a topic on which they have specific expertise compared to when informally answering a common-knowledge question; (b) whether listeners judge the voices of these men and women as more competent and authoritative when giving expert advice than in the control condition; and (c) whether the production or perception of voice modulation differs for male versus female speakers.

Study 1: Acoustic Correlates of Authoritative Speech

Method and Materials

Participants

Fifty-one professional men and women took part in the study as speakers (27 men and 24 women aged 30–55 years). All participants were academic faculty members and scientists with at least a PhD degree, recruited from various universities in Wrocław, Poland.

Procedure

Two trained research assistants (a man and a woman, unknown to participants) visited the faculty individually during their office hours and introduced themselves as employees of a radio broadcasting agency. These research assistants then engaged the faculty member in two types of speech conditions: (1) *control speech*, wherein the subjects were asked to provide directions to the administrative offices (in Polish, “dziekanat”) on that given campus, which is common knowledge to all university faculty; and (2) *authority speech*, wherein the research assistants indicated that they ran a radio program for young scholars titled, “How to become a scientist, and is it worth it?”, and invited the faculty member to express her or his opinion on this topic. In the authority speech condition, the research assistants repeatedly highlighted the expertise of the faculty member (e.g., “*Dr. [last name], you are an expert and authority in this area, please tell us...*”). In the control speech condition, all participants knew where the administrative offices were located and provided accurate directions without hesitation. Condition order was counter-balanced across participants. All voice recordings were taken in the faculty members’ native language (Polish).

The voices of participants were recorded during both speech conditions using a Zoom H4n digital audio recorder with built-in cardioid condenser microphone at a distance of approximately 40–50 cm from the participant. Audio recordings were made in mono at a sampling rate of 44.1 kHz and 16-bit quantization, and then uploaded to a computer as uncompressed WAV files. These procedures were approved by the Institute of Psychology’s Ethics Review Board at the University of Wrocław. All participants provided informed consent for their voice recordings to be acoustically analyzed for the purposes of this study and also to be broadcast on the radio.¹

Voice Measurement

Acoustic measurements were performed in Praat v.6.0.17 (Boersma and Weenink 2016). For each voice recording (two speech conditions per participant), we measured mean fundamental frequency (F_0), minimum and maximum F_0 , the standard deviation of F_0 (F_0 SD) and its range (F_0 range), formant position (P_f), formant spacing (ΔF), and apparent vocal tract length ($VTL(\Delta F)$). All measures were taken from voiced speech segments only. Fundamental frequency measures were taken in Hz and also converted into equivalent rectangular bandwidths (ERBs), a semi-logarithmic scale that controls for the decrease in perceptual

¹ Currently a real radio program titled “How to become a scientist, and is it worth it?” is being prepared for the local radio station, and will feature the advice given by our faculty participants.

salience (perceived pitch) of F_0 as frequency increases. The control speech (duration: estimated marginal mean, $EMM=10.1$, $SD=6.7$ s) and authority speech (duration: $EMM=59.7$, $SD=23.5$ s) recordings were analyzed in full, and recording duration was controlled in statistical models (see "Study 1 Results" section).

All F_0 parameters were measured using Praat's autocorrelation algorithm with a search range set to 75–300 Hz for men and 100–500 Hz for women. Formants F_1 – F_4 were measured using Praat's Burg linear predictive coding algorithm with the initial settings of maximum formant set to 5000 Hz for men and 5500 Hz for women. Formant frequencies F_1 – F_4 were measured at each glottal pulse (as in Puts et al. 2012a). We utilized the median of these values across the sound file as the median is less influenced than the mean by errors made by Praat in identifying formants (e.g., F_2 misattributed as F_1). The fundamental frequency and formant measures obtained agree well with weighted population-level averages (Pisanski et al. 2014).

From the mean F_1 – F_4 values we computed P_f (Puts et al. 2012a), the average standardized formant value for each subject, following the equation:

$$P_f = \frac{\sum_{i=1}^n F'_i}{n}$$

where F'_i is the standardized i th formant, and n is the number of measured formants. So as not to bias the mean and SD formant measures used in standardizing toward the more numerous sex in the sample, we used the unweighted average of the women and men's means as the overall mean, and used the following equation for the pooled SD :

$$S_p = \sqrt{\frac{(n1 - 1)S_1^2 + (n2 - 1)S_2^2}{n1 + n2 - 2}}$$

We also computed formant spacing, ΔF , a measure of the distance among adjacent formants (Reby and McComb 2003). Formant spacing effectively predicts vocal tract length, body size and shape of both men and women (i.e., taller and larger individuals have more closely spaced formants and longer vocal tracts, Pisanski et al. 2016b). Each formant is related to ΔF by the following equation:

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

where i represents the specific formant. To compute ΔF , we plotted mean formant values for each subject against $(2i - 1)/2$, the expected increments of formant spacing as predicted by Reby and McComb's (2003) vocal tract model, where ΔF is equal to the slope of the linear regression line with an intercept of 0. From this, we estimated the apparent vocal tract length of each subject in cm using the following equation:

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

where c is 35 000 cm/s, the approximate speed of sound in the human vocal tract (for additional details see Reby and McComb 2003; see also Table 1).

Table 1 MANOVA model with speech condition (control vs. authority), sex of speaker (male vs. female) and condition order as variables

		Mean F_0	Min F_0	Max F_0	F_0 SD	F_0 range	P_f	ΔF	VTL (ΔF)
Between-subject effects									
Sex	F	211.27**	103.3**	276.33**	47.74**	169.43**	68.52**	403.41**	369.36**
	η^2	0.82	0.69	0.86	0.50	0.78	0.59	0.90	0.89
Order	F	0.02	0.18	0.23	0.15	0.12	1.15	0.66	0.45
	η^2	<0.01	0.01	<0.01	<0.01	<0.01	0.03	0.01	0.01
Sex $\times$ order	F	1.33	0.31	0.02	0.90	<0.01	1.66	0.13	0.09
	η^2	0.03	<0.01	<0.01	0.02	<0.01	0.03	<0.01	<0.01
Within-subject									
Condition	F	6.76*	1.40	0.93	3.64	1.84	7.82*	1.92	2.18
	η^2	0.13	0.03	0.02	0.06	0.04	0.14	0.04	0.05
Condition $\times$ sex	F	9.21**	2.35	0.81	1.55	1.97	0.95	0.13	0.33
	η^2	0.16	0.05	0.02	0.03	0.04	<0.01	<0.01	<0.01
Condition $\times$ order	F	0.11	0.09	3.73	0.22	4.13*	0.97	0.69	0.36
	η^2	<0.01	<0.01	0.07	<0.01	0.08	0.02	0.01	0.02
Condition $\times$ sex $\times$ order	F	0.12	0.31	1.74	0.08	2.26	0.97	0.05	0.62
	η^2	<0.01	<0.01	0.04	<0.01	0.05	0.02	0.01	0.01
Covariate									
Speech duration difference	F	0.06	3.81	0.46	0.22	1.41	2.54	0.12	0.09
	η^2	<0.01	0.08	0.01	<0.01	0.03	0.05	<0.01	<0.01
Speech duration difference $\times$ condition	F	0.36	0.39	0.37	0.56	0.66	0.55	0.15	0.20
	η^2	0.04	<0.01	<0.01	0.01	0.01	0.01	<0.01	<0.01

** $p < .01$; * $p < .05$

Fig. 1 Estimated marginal means for all voice parameters between the authority and control speech conditions ($\pm$ standard error mean). *Note* P_f values were multiplied by two orders of magnitude ($*10^2$) and ΔF values were divided by an order of magnitude ($*10^{-1}$) to allow all voice parameters to be plotted on a comparable scale. P_f values were negative for men and are plotted here as absolute values, again for comparative purposes. Original raw values are given as labels next to each bar

Study 1: Results

To test for differences in voice parameters between speech conditions, we conducted an omnibus repeated measures MANOVA with *speech condition* (control vs. authority) as a within-subject variable, and *sex of speaker* (male vs. female) and *condition order* as between-subject variables. Because the average length of voice recordings collected in the authority speech condition was significantly longer ($M = 57.2 \pm .09$) than the average length of control speech recordings ($M = 10.3 \pm 3.2$; $t_{51} = 15.4$, $p < .001$), the difference in recording duration between conditions (authority-control) was computed for each participant and included as a covariate in the model.

We observed main effects of condition on mean F_0 and P_f as well as expected main effects of speaker sex on all voice parameters (see Table 1). Pairwise comparisons confirmed higher values for all voice parameters among women compared to men, and longer VTL(ΔF) estimates for men than women, as expected due to sexual dimorphism in the human larynx and vocal tract (Titze 1989). Critically, for both sexes, authority speech was characterized by lower mean F_0 (voice pitch) and lower P_f (formant position) compared to control speech. Furthermore, the model revealed a significant interaction between condition and sex on mean F_0 , indicating that women lowered their voice pitch more than did men in the authority speech condition (see Fig. 1). We also found an interaction between speech condition and order on F_0 range, indicating that when subjects partook first in the authority condition followed by the control condition, F_0 range was lower than for the opposite order of conditions. Speech duration showed no main or interaction effects (Table 1).

Study 2: Listeners' Voice-Based Ratings of Authority and Competence

Method

Participants

Thirty-nine participants (31 males) aged 20–49 years ($M = 33.5 \pm 7.6$) completed the playback experiment. Participants were recruited via online advertisements calling for foreign-speaking adults currently staying in Wrocław, Poland. Participants represented a wide range of nationalities (Bulgaria $n = 1$, Canada $n = 2$, Croatia $n = 1$, Finland $n = 2$, France $n = 2$, Germany $n = 4$, Greece $n = 3$, Iraq $n = 1$, Italy $n = 4$, Korea $n = 1$, Luxembourg $n = 1$, Spain $n = 8$, Turkey $n = 1$, UK $n = 4$, USA $n = 3$, unspecified $n = 1$). The inclusion criteria were that participants (1) could not understand Polish, so that the linguistic content of speech samples would not influence their ratings, and (2) could understand written English, as this was the language in which recruitment and experimental materials were presented.

Voice Stimuli

Voice stimuli were derived from Study 1 and thus included voice clips taken from the authority and control speech conditions of each recorded faculty member. To prepare voice recordings for playback, segments of multi-voicing, acute noise, and nonverbal vocalizations (e.g., laughter) were first manually removed in Praat v. 6.0.29 (Boersma and Weenink 2016). Due to a significant difference in duration between speech conditions (see Study 1), stimuli were then standardized for recording length by extracting the first 5 s of each voice recording for playback. Voice stimuli were amplitude normalized to 70 dB RMS SPL and played back at constant amplitude within participants.

Procedure

Participants completed the playback experiment privately in a silent room via a custom controlled interface designed in Visual Studio Code (v.1.4) to present voice stimuli and archive participants' responses. Voice stimuli were played through Philips SHE4205WT/00 professional headphones at a constant preset volume. Stimuli were paired by condition and within speakers, such that each voice pair comprised authority and control speech from the same speaker, with pairing order randomized. Participants were instructed that they would hear pairs of voices, and on each trial, would be asked to choose which of the two voices sounded more 'competent', or in a separate question, which of the two voices sounded more 'authoritative'. A *competent* person was defined as, "a person having the necessary knowledge or skill to do something successfully or efficiently" (Oxford Online Dictionary 2018a). A person with *authority* was defined as, "a person with extensive or specialized knowledge about a subject; an expert" (Oxford Online Dictionary 2018b).

Study 2: Results

Linear Mixed Models with speaker sex included as a fixed factor indicated that listeners chose voice stimuli from the authority speech condition as relatively more competent than control speech stimuli on 62% of trials, with no effect of speaker sex ($F_{1,206} = .004, p = .95$).

Fig. 2 Percentage of stimuli in voice pairs (authority vs. control speech from the same speaker) chosen by listeners as more competent (left 3 bars) or more authoritative (right 3 bars). Two-tailed one-sample t tests against chance (50%) where *** $p < .001$; ** $p < .01$

Likewise, listeners chose authority speech stimuli as more authoritative than control speech stimuli on 57% of trials. However, the effect of voice condition on listeners' judgments of authority differed for male and female voices ($F_{1,2026} = 5.16$, $p = .023$), indicating that authoritative speech stimuli were chosen as more authoritative than control speech stimuli more often for male ($M = 0.59 \pm .015$) than for female speakers ($M = 0.54 \pm .016$). These effects were corroborated through a series of two-tailed t tests, confirming that despite this sex difference, authority speech stimuli were judged as relatively more authoritative, and more competent, at above-chance levels for both male ($t_{1052} = 6.18$, $p < .001$; $t_{1052} = 7.96$, $p < .001$) and female speakers ($t_{974} = 2.73$, $p = .006$; $t_{974} = 7.75$, $p < .001$; see Fig. 2).

Discussion

Voice researchers have only recently begun to investigate whether people modulate non-verbal parameters of their voices in various social contexts, and this research has focused almost exclusively on mating contexts (Fraccaro et al. 2011; Hughes et al. 2010, 2014; Leongómez et al. 2014; Pisanski et al. 2018; Puts et al. 2006). Our results demonstrate that male and female professionals (scientists working as faculty at various universities) lower their vocal fundamental and formant frequencies when asked to provide their expert opinion compared to when giving easy campus directions. While both sexes lowered their mean F_0 (voice pitch) in an authoritative context, women lowered their pitch significantly more than did men. Additionally, playback experiments showed that listeners judged the voices of these men and women as more authoritative and more competent when giving expert advice. The social judgments of these foreign-speaking listeners were necessarily based on non-linguistic parameters of the speakers' voices. Our results therefore support the prediction that people modulate the nonverbal parameters of their voices in social contexts to elicit appropriate social appraisals. This is line with *between* individual studies, showing that low fundamental and formant frequencies predict judgments of dominance, competence, intelligence and trustworthiness, wherein such social traits are typically favored in professional social contexts (Klofstad et al. 2012, 2015; Kreiman and Sidtis 2011; Puts et al. 2012b; Tigue et al. 2012).

The finding that women lowered their mean voice pitch more than did men in an authoritative context (on both a physical, F_0 , and perceptual, ERB scale) may reflect both

anatomical and cultural factors. Due to anatomical and hormonal influences, women's voice pitch is on average almost twice as high as men's (Titze 1989). Hence, women may need to lower their voice pitch by a greater absolute amount compared to men to achieve similarly salient effects. Sex differences have also been observed in volitional voice modulation tasks (vocal modulation of attractiveness: Hughes et al. 2014; vocal modulation of body size: Pisanski et al. 2016c) and in studies examining voice modulation in real life settings (Pisanski et al. 2018). In addition, men may habitually speak closer to the bottom of their pitch range, reducing the amount by which they are able to lower their pitch contextually. At the same time, women appear aware of vocal stereotypes linked to masculinity, and may actively modulate various voice parameters to assume a more masculine presentation in professional contexts (Anderson et al. 2014; Karpf 2006; Klofstad et al. 2012). Interestingly, the results of our playback study showed that, while both men and women were judged more often as authoritative and competent when giving expert advice, this difference was more pronounced for authority judgments of male speakers. Taken together this suggests that while women may modulate their voice pitch more than men in an authoritative context, the advantage gained by such voice modulation may still be greater for men. However, as the goal of the playback study was to examine the effect of naturally modulated speech on listeners' social judgments, we did not systematically manipulate voice parameters, and thus the relative contribution of pitch and formants (or other voice parameters) to listeners' judgments cannot be ascertained.

Indeed, a great deal more research is needed to better understand which vocal parameters men and women modulate, under which social conditions, and how these modulations influence listeners' judgments. The present research highlights clear sex differences in the production and perception of nonverbal vocal parameters in modulated speech, and this is an important avenue for future work. Future studies may also test whether age influences the degree to which men or women modulate their voices in a professional context, where, for example, younger individuals may lower their voice pitch relatively more than older individuals to project authority. Moreover, although we found that men and women spoke with lower mean pitch (mean F_0) and lower formant position (P_f) when giving expert advice, we found no differences in pitch contour and variability or in formant-based estimates of vocal tract length between speech conditions after controlling for speech duration. As pitch variability and formant spacing can offer socially-relevant information related to, for example, masculinity and formidability (Pisanski et al. 2014; Puts et al. 2012a), follow-up studies should examine if speakers modulate these voice parameters in other social contexts. Finally, while statistically significant, the overall difference in listeners' judgments of competence or authority between speech conditions was small. How various vocal parameters or social variables contribute to the effective manipulation of listeners' judgments in voice modulation remains largely unknown.

Laboratory studies examining controlled “on demand” voice modulation have shown that both men and women lower their voice pitch and formants to sound physically larger (Pisanski et al. 2016c) or more masculine (Cartei et al. 2012). However, Hughes et al. (2014) prompted participants to “sound confident, intelligent, or dominant”, and found that both men and women raised their voice pitch in each of these contexts by 5–13 Hz. The authors also found that men raised their voice pitch to “sound attractive”, whereas women lowered their voice pitch in this condition. These findings generally oppose the patterns predicted and supported by a large body of previous work (reviewed in Feinberg 2008; Pisanski and Feinberg 2019; Puts et al. 2012b), including laboratory studies of voice modulation utilizing more ecologically valid procedures (Puts et al. 2006), though a recent speed dating study found that women also lowered their voice pitch toward desirable potential

mates whom they showed a personal preference for (Pisanski et al. 2018). Thus, the results of past research on social voice modulation are mixed, and although studying vocal modulation in real world social scenarios has the limitation of low experimental control, such field research boasts high ecological validity, and together both approaches (lab and field) can provide novel insights for the thriving field of nonverbal communication, helping to elucidate the various factors contributing to social voice modulation.

Funding This research was made possible by funding from the Polish Ministry of Science and Higher Education (Grant Juventus Plus, IP2012 046672). AO was supported by the Ministry of Science and Higher Education (#626/STYP/12/2017). KP was supported by the Ministry of Science and Higher Education (#530/STYP/10/2015) and the Foundation for Polish Science (START Program, 2015–2018).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided 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.

References

- Anderson, R. C., Klostad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE*, 9(5), e97506.
- Babel, M., McGuire, G., & King, J. (2014). Towards a more nuanced view of vocal attractiveness. *PLoS ONE*, 9(2), e88616. <https://doi.org/10.1371/journal.pone.0088616>.
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer v 6.0.21*. [computer software]
- Cartei, V., Cowles, H. W., & Reby, D. (2012). Spontaneous voice gender imitation abilities in adult speakers. *PLoS ONE*, 7(2), e31353. <https://doi.org/10.1371/journal.pone.0031353>.
- Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica*, 16(1), 23–28.
- Feinberg, D. R. (2008). Are human faces and voices ornaments signaling common underlying cues to mate value? *Evolutionary Anthropology: Issues, News, and Reviews*, 17(2), 112–118.
- Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., et al. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology*, 9(1), 57–67. <https://doi.org/10.1556/JEP.9.2011.33.1>.
- Gregory, S. W., Jr., & Gallagher, T. J. (2002). Spectral analysis of candidates' nonverbal vocal communication: Predicting US presidential election outcomes. *Social Psychology Quarterly*, 65(3), 298–308.
- Gregory, S. W., Webster, S., & Huang, G. (1993). Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language & Communication*, 13(3), 195–217.
- Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nature*, 21(4), 406–427. <https://doi.org/10.1007/s12110-010-9101-5>.
- Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and physiological changes in response to the physical attractiveness of conversational partners. *Journal of Nonverbal Behavior*, 34(3), 155–167.
- Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior*, 38(1), 107–127. <https://doi.org/10.1007/s10919-013-0163-z>.
- Hunter, E. J., & Titze, I. R. (2010). Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. *Journal of Speech, Language, and Hearing Research*, 53(4), 862–875. [https://doi.org/10.1044/1092-4388\(2009/09-0040\)](https://doi.org/10.1044/1092-4388(2009/09-0040)).
- Karpf, A. (2006). *The human voice*. New York: Bloomsbury.
- Klostad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. *PLoS ONE*, 10(8), e0133779. <https://doi.org/10.1371/journal.pone.0133779>.

- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society of London B: Biological Sciences*. <https://doi.org/10.1098/rspb.2012.0311>.
- Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. New Jersey: Wiley-Blackwell.
- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., et al. (2014). Vocal modulation during courtship increases perceptivity even in naive listeners. *Evolution and Human Behavior*, 35(6), 489–496. <https://doi.org/10.1016/j.evolhumbehav.2014.06.008>.
- Leongómez, J. D., Mileva, V. R., Little, A. C., & Roberts, S. C. (2017). Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PLoS ONE*, 12(6), e0179407.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111(981), 855–869.
- Oxford Online Dictionary. (2018a). Authority. Retrieved April 1, 2018, from <https://en.oxforddictionaries.com/definition/authority>.
- Oxford Online Dictionary. (2018b). Competent. Retrieved April 1, 2018, from <https://en.oxforddictionaries.com/definition/competent>.
- Pisanski, K., & Bryant, G. A. (2019). The evolution of voice perception. In N. S. Eidsheim & K. L. Meizel (Eds.), *The Oxford handbook of voice studies* (pp. 269–300). New York, NY: Oxford University Press.
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016a). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, 20(4), 304–318.
- Pisanski, K., & Feinberg, D. R. (2019). Vocal attractiveness. In S. Frühholz & P. Belin (Eds.), *Oxford handbook of voice perception* (pp. 607–625). New York, NY: Oxford University Press.
- Pisanski, K., Fraccaro, P. J., Tigues, C. C., O'Connor, J. J. M., Röder, S., Andrews, P. W., et al. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, 95, 89–99. <https://doi.org/10.1016/j.anbehav.2014.06.011>.
- Pisanski, K., Jones, B. C., Fink, B., O'Connor, J. J. M., DeBruine, L., Roder, S., et al. (2016b). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, 112, 13–22. <https://doi.org/10.1016/j.anbehav.2015.11.008>.
- Pisanski, K., Mora, E., Pisanski, A., Reby, D., Sorokowski, P., Franckowiak, T., et al. (2016c). Volitional exaggeration of body size through fundamental and formant frequency modulation in humans. *Scientific Reports*, 6, 34389. <https://doi.org/10.1038/srep34389>.
- Pisanski, K., Nowak, J., & Sorokowski, P. (2016d). Individual differences in cortisol stress response predict increases in voice pitch during exam stress. *Physiology & Behavior*, 163, 234–238. <https://doi.org/10.1016/j.physbeh.2016.05.018>.
- Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., & Reby, D. (2018). Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B*, 285(1893), 20181634. <https://doi.org/10.1098/rspb.2018.1634>.
- Puts, D. A. (2016). Human sexual selection. *Current Opinion in Psychology*, 7, 28–32. <https://doi.org/10.1016/j.copsyc.2015.07.011>.
- Puts, D., Apicella, C. L., & Cardenas, R. A. (2012a). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 601–609. <https://doi.org/10.1098/rspb.2011.0829>.
- Puts, D., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283–296. <https://doi.org/10.1016/j.evolhumbehav.2005.11.003>.
- Puts, D., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., et al. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152830.
- Puts, D., Jones, B. C., & DeBruine, L. M. (2012b). Sexual selection on human faces and voices. *The Journal of Sex Research*, 49(2–3), 227–243. <https://doi.org/10.1080/00224499.2012.658924>.
- Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, 65(3), 519–530. <https://doi.org/10.1006/anbe.2003.2078>.
- Scherer, K. (2019). Acoustic patterning of emotion vocalizations. In S. Frühholz & P. Belin (Eds.), *Oxford handbook of voice perception* (pp. 61–91). New York, NY: Oxford University Press.
- Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological Science*, 26(6), 877–891. <https://doi.org/10.1177/0956797615572906>.

- Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research: Advances in vocal communication research. *Journal of Zoology*, 280(3), 221–236. <https://doi.org/10.1111/j.1469-7998.2009.00661.x>.
- Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216. <https://doi.org/10.1016/j.evolhumbehav.2011.09.004>.
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, 85(4), 1699–1707. <https://doi.org/10.1121/1.397959>.
- Tuomi, S. K., & Fisher, J. E. (1979). Characteristics of simulated sexy voice. *Folia Phoniat*, 31, 242–249.
- Wang, Y., Houston, D. M., & Seidl, A. (2019). Acoustic properties of infant-directed speech. In S. Frühholz & P. Belin (Eds.), *Oxford handbook of voice perception* (pp. 93–115). New York, NY: Oxford University Press.

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

Affiliations

Piotr Sorokowski¹ · David Puts^{2,3,4} · Janie Johnson² · Olga Żółkiewicz¹ · Anna Oleszkiewicz^{1,5} · Agnieszka Sorokowska^{1,6} · Marta Kowal¹ · Barbara Borkowska⁷ · Katarzyna Pisanski^{1,8}

¹ Institute of Psychology, University of Wrocław, Wrocław, Poland

² Department of Anthropology, The Pennsylvania State University, University Park, USA

³ Center for Brain, Behavior, and Cognition, The Pennsylvania State University, University Park, USA

⁴ Center for Human Evolution and Diversity, The Pennsylvania State University, University Park, USA

⁵ Smell and Taste Center, Department of Otorhinolaryngology, TU Dresden, Dresden, Germany

⁶ Department of Psychotherapy and Psychosomatic Medicine, TU Dresden, Dresden, Germany

⁷ Department of Human Biology, University of Wrocław, Wrocław, Poland

⁸ Present Address: Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS, UMR 9197, University of Lyon/Saint-Etienne, Saint-Etienne, France