

Low Vocal Pitch Preference Drives First Impressions Irrespective of Context in Male Voices but Not in Female Voices



Perception

2016, Vol. 45(8) 946–963

© The Author(s) 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0301006616643675

pec.sagepub.com



Maria S. Tsantani

School of Psychology, University of Glasgow, Glasgow, UK

Pascal Belin

Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK; Institut des Neurosciences de la Timone, CNRS, Université Aix-Marseille, Marseille, France; Department of Psychology, International Laboratories for Brain, Music and Sound, Université de Montréal, McGill University, Montreal, QC, Canada

Helena M. Paterson and Phil McAleer

School of Psychology, University of Glasgow, Glasgow, UK

Abstract

Vocal pitch has been found to influence judgments of perceived trustworthiness and dominance from a novel voice. However, the majority of findings arise from using only male voices and in context-specific scenarios. In two experiments, we first explore the influence of average vocal pitch on first-impression judgments of perceived trustworthiness and dominance, before establishing the existence of an overall preference for high or low pitch across genders. In Experiment 1, pairs of high- and low-pitched temporally reversed recordings of male and female vocal utterances were presented in a two-alternative forced-choice task. Results revealed a tendency to select the low-pitched voice over the high-pitched voice as more trustworthy, for both genders, and more dominant, for male voices only. Experiment 2 tested an overall preference for low-pitched voices, and whether judgments were modulated by speech content, using forward and reversed speech to manipulate context. Results revealed an overall preference for low pitch, irrespective of direction of speech, in male voices only. No such overall preference was found for female voices. We propose that an overall preference for low pitch is a default prior in male voices irrespective of context, whereas pitch preferences in female voices are more context- and situation-dependent. The present study confirms the important role of vocal pitch in the formation of first-impression personality judgments and advances understanding of the impact of context on pitch preferences across genders.

Corresponding author:

Phil McAleer, School of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, UK.

Email: philip.mcaleer@glasgow.ac.uk

Keywords

social perception, social interaction, vocal pitch, dominance, trustworthiness, first impressions

When forming first impressions of others' personalities in everyday social life, people take into account information from multiple senses: they see an "attractive" face, hear a "trustworthy" voice, and feel a "dominant" handshake. Such judgments are often formed rapidly in order to direct a decision about the appropriate course of action toward the person in question, and, together with the face, the human voice is one of the most important providers of information utilized when forming these first impressions. Indeed, Belin, Bestelmeyer, Latinus, and Watson (2011) propose that the human voice acts as an "auditory face," allowing for the rapid extraction of paralinguistic socially relevant information, including information about personality (Belin et al., 2011). Recent work has highlighted such personality judgments, with McAleer, Todorov, and Belin (2014) showing that people can consistently establish first-impression judgments from voices after listening to only subsecond extracts.

Within the first-impression personality judgment literature, studies using both faces and voices have indicated that trustworthiness and dominance are the most prominent of perceived traits (e.g., McAleer et al., 2014; Oosterhof & Todorov, 2008; Sutherland et al., 2013). Oosterhof and Todorov (2008) proposed that the adaptive mechanisms for detecting threat from others have become overgeneralized from attributing momentary emotional states to static personality traits, with the authors linking the percept of trustworthiness to the intent of threat, and the percept of dominance to the ability to carry out the threat. Thus, the overgeneralization of subtle facial cues leads to very rapid, but not necessarily accurate, first-impression judgments of trustworthiness (i.e., whether a person should be approached or avoided) and dominance (i.e., whether a person is physically weak or strong; McArthur & Baron, 1983; Oosterhof & Todorov, 2008; Secord, 1958; Zebrowitz & Montepare, 2008). For similar purposes, a recent study by McAleer et al. (2014) suggested a parallel functionality of personality judgments in voices, and showed that, as with faces, numerous vocal personality traits can be represented in a two-dimensional "social voice space" in which valence (i.e., trustworthiness and likeability) forms the primary dimension and dominance forms the second dimension.

Such estimates of personality from the voice are found to be driven by differences in vocal acoustics between individuals (e.g., Apicella & Feinberg, 2009; Ferdenzi et al., 2013; McAleer et al., 2014; Scherer, 1972; Zuckerman & Driver, 1989). A key acoustical feature is the fundamental frequency of the voice (f_0), that is, the rate of vocal fold vibration (Fitch, 2000), otherwise called "pitch." Vocal pitch is a highly sexually dimorphic vocal characteristic (Puts, Jones, & DeBruine, 2012). Up until adolescence, vocal pitch shows a steady decrease in both genders; in adolescence, however, it decreases much faster in males than in females, resulting in adult male voices having fundamental frequencies that are, on average, twice as low as those of adult female voices (male range: 80 Hz to 185 Hz; female range: 165 Hz to 255 Hz; Titze, 1989). Hence, males tend to have deeper voices (Huber, Stathopoulos, Curione, Ash, & Johnson, 1999; Puts, Apicella, & Cárdenas, 2011). Numerous studies have highlighted the importance of pitch as one of the main vocal signals in determining various aspects about a speaker, such as age, gender, and emotion (cf. Latinus & Belin, 2011), as well as its importance to our interactions and decisions (Apicella, Feinberg, & Marlowe, 2007; Apicella & Feinberg, 2009; Klofstad, Anderson, & Peters, 2012; Nass & Lee, 2001). Pitch may be a key component in judging certain personality traits from voices due to its malleability. An analogous relationship has already been

proposed for the importance of the eyes and mouth in establishing personality and emotion in faces (McAlear et al., 2014; Oosterhof & Todorov, 2008).

Addressing the key traits individually, variations in vocal pitch between individuals have previously been shown to be negatively related to androgen levels (Dabbs Jr. & Mallinger, 1999; Evans, Neave, Wakelin, & Hamilton, 2008; Newman, Butler, Hammond, & Gray, 2000), and as such, dominance has consistently been associated with low vocal pitch in men—whether the judgment be about attracting a mate, selecting a political leader, or adopting the appropriate voice for business (Feinberg et al., 2006; Jones, Feinberg, DeBruine, Little, & Vukovic, 2010; Klofstad et al., 2012; McAlear et al., 2014; Ohala, 1982; Puts, Gaulin, & Verdolini, 2006; Puts, Hodges, Cárdenas, & Gaulin, 2007; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012; Vukovic et al., 2011; Watkins et al., 2010; Wolff & Puts, 2010). However, in contrast, the relationship between vocal pitch and dominance in females has received less attention in the literature. Ohala (1982) proposed that dominance is associated with low vocal pitch regardless of gender. Likewise, Jones et al. (2010) agreed that low-pitched female voices were perceived as more dominant compared with high-pitched voices; a finding supported by Borkowska and Pawlowski (2011). However, McAlear et al. (2014), utilizing a Likert rating on a varied stimuli set (as opposed to the two-alternative forced-choice [2AFC] task that was used in the majority of the previously mentioned studies) found that dominance was more associated with high vocal pitch than with low vocal pitch in females. Ultimately, the relationship in female voices is understudied and findings are inconsistent.

In regard to perceived trustworthiness, as with dominance in females, few studies have specifically targeted the effect of vocal pitch on this trait, with results proving largely inconsistent. In male voices, two studies exploring the effect of vocal pitch on election voting and personality traits found contrasting results; Tigue et al. (2012) found that low-pitched candidates were perceived as more trustworthy, whereas Klofstad et al. (2012) found no significant preference in pitch with regard to trustworthiness. Similarly, exploring the effect of seeking short- versus long-term relationships, Vukovic et al. (2011) showed no significant preference in male pitch when female participants judged the voices for trustworthiness. However, McAlear et al. (2014) found that high-pitched male voices were rated as more trustworthy compared with low-pitched voices. Considering trustworthiness in female voices, as is the case with dominance, there is a lack of studies, and again a lack of consistent findings. Klofstad et al. (2012) found that low-pitched female speakers were judged as being more trustworthy in a political scenario, while, in contrast, McAlear et al. (2014) found that trustworthiness in female voices was not influenced by average voice pitch but by the vocal glide and intonation; how the pitch moves. In short, in both genders, there is a lack of studies examining the role of pitch and the preference of high versus low pitch in trustworthiness judgments, with previous findings having failed to provide a consistent answer.

One source of variability between the mentioned studies, and perhaps an explanation for the diverse findings, is the utterance selected for the voice stimuli. The majority of studies exploring judgments of personality in voices have done so within a specified context, such as the election of political leaders (Klofstad et al., 2012; Tigue et al., 2012) and relationship preferences (Vukovic et al., 2011), utilizing either content-relevant voice stimuli, or, conversely, short nonsocially relevant stimuli such as vowel sounds (e.g., Jones et al., 2010). Thus, given that it has been suggested that preferences for low or high vocal pitch can be moderated by social context (Jones, Feinberg, DeBruine, Little, & Vukovic, 2008; Vukovic et al., 2008), the generalizability of these findings across different, noncontextual, or ambiguous contexts becomes limited. Moreover, the majority of studies exploring vocal pitch

in relation to both dominance and trustworthiness have focused on male voices and, therefore, it is unclear how well previous findings can be generalized to female voices.

Thus, the first aim of the study was to explore the effect of vocal pitch on first-impression judgments of both dominance and trustworthiness, in both male and female voices, using noncontextual, yet socially relevant, stimuli. High- and low-pitched versions of male and female voice recordings of the word “hello” (McAleer et al., 2014) were created by raising and lowering the fundamental frequency of original recordings by approximately 20 Hz (Klofstad et al., 2012). The resulting voice samples were individually temporally reversed so as to minimize possible influences of semantic content on participants’ preferences but to maintain acoustical information and to enable listeners to recognize them as voices (Fleming, Giordano, Caldara, & Belin, 2014). It has been demonstrated that although reversed speech stimuli are unintelligible they activate the same brain regions as natural speech (Binder et al., 2000). It is proposed that the use of reversed words as stimuli maintains the social nature of language, in contrast to the use of vowel sounds, and, at the same time, the use of subsecond voice recordings ensures that the time duration is sufficient to produce first-impression judgments (McAleer et al., 2014). The reversal of speech further counters the argument that the McAleer et al. (2014) stimuli were merely ambiguous and not noncontextual; temporally reversed speech moves the stimuli, we argue, closer to being completely noncontextual given that a listener cannot understand what is being said or who is being addressed. In short, the stimuli are socially relevant in that people perceive them as words, though they may not understand what is being said.

Based on the perceived association between low vocal pitch and aggressive potential in men (Bartholomew & Collias, 1962; Puts et al., 2012) and with characteristics of social dominance such as assertiveness, charisma, and leadership skills in both men and women (Bolinger, 1964; Borkowska & Pawlowski, 2011; Ohala, 1983; Puts et al., 2007), it was predicted that low-pitched male and female voices would be perceived as more dominant, compared with the high-pitched voices. Second, in accordance with McAleer et al.’s (2014) finding that high vocal pitch is associated with higher perceived trustworthiness in male voices, it was expected that high-pitched male voices would be perceived as more trustworthy compared with the low-pitched voices. Finally, in line with Klofstad et al. (2012), it was expected that the low-pitched female voices would be perceived as more trustworthy compared with their high-pitched counterparts.

Experiment I

Methods

Ethics statement. All procedures for this experiment were approved by the University of Glasgow ethics committee for the School of Psychology, in accordance with the 1964 Declaration of Helsinki. Before taking part in the experiment, participants provided written consent for their participation after reading a form reminding them of their freedom to withdraw at any point, and of the anonymity and confidentiality of the provided data.

Participants. A total of 40 participants (13 male; average age = 24 years; $SD = 3$ years) from the University of Glasgow subject pool took part in this experiment. No monetary incentive was given for taking part; partial course credit was given where relevant.

Stimuli and apparatus. The voice recordings were obtained from an existing sample of 64 (32 male) original voice recordings of the word “hello,” with predetermined ratings of dominance

and trustworthiness (McAleer et al., 2014). A total of 20 (10 female and 10 male) voice recordings were selected from the middle range of the ratings for dominance and 20 (10 female and 10 male) voice recordings were selected from the middle range of ratings for trustworthiness: The original pitch values for the chosen stimuli were taken to be within an accepted range for male and female pitches (Titze, 1989). This selective method was used in order to minimize the possibility of the original voices containing characteristics, besides vocal pitch, that rendered them as extremely dominant or trustworthy, or extremely nondominant or untrustworthy; due to this restrictive range, five female voices and two male voices appeared in both the dominant and trustworthy sets of voices. The PRAAT phonetic analysis programme (v. 5.1.25; Boersma, 2001) was used to create a high- and a low-pitched version of each recording by altering the frequency approximately ± 20 Hz: Due to fluctuations in PRAAT, average pitch shift was 19.74 Hz in male voices and 19.21 Hz in female voices. A difference of approximately 40 Hz between voices has been shown to be large enough to cause significant differences in perceptual judgments of vocal personality traits (e.g., Borkowska & Pawlowski, 2011; Feinberg, Jones, Little, Burt, & Perrett, 2005). All voice samples were temporally reversed to minimize semantic content but to maintain acoustical information and to enable listeners to recognize them as voices (Fleming et al., 2014). Finally, the stimuli were normalized for power (Root Mean Square (RMS)) and loudness via MATLAB (the Mathworks). The average duration of the stimuli across the whole experiment was approximately 400 ms. A total of 40 pairs of voice samples (80 individual samples) were created using this process: high and low pitch for 10 male and 10 female from the trustworthiness scale; high- and low-pitch samples for 10 male and 10 female from the dominance scale. The experiment was presented on E-Prime v2.0 software running on Dell Inspiron One 2320 (Intel Core i5) PCs. All relevant values for the stimuli are reported in Table S1 for female voices and Table S2 for male voices (see Supplementary Information).

Procedure. The experiment took place in the experimental laboratories of the University of Glasgow. Participants were required to complete a 2AFC task during which they listened to pairs of voices comprising high- and low-pitched versions of the original recordings. The sound samples were presented through headphones (participants' own) connected to a computer with the sound set at approximately 80 dB Sound Pressure Level (SPL): System volume was measured prior to the experiment using a standard headphone set (Sennheiser Beyerdynamic DT 770 PRO 250 OHM) and sound meter. The 2AFC task has been used previously to measure the effect of vocal pitch on personality judgments (e.g., Feinberg et al., 2005). That choices are made between pairs of the same voice, as opposed to different voices, eliminates the possibility that other characteristics besides vocal pitch cause participants to choose one voice over the other: for example, glide, intonation, and Harmonic-to-Noise ratio (HNR) which have all previously been shown to have influence on judgments (McAleer et al., 2014).

At the beginning of the experiment, participants were informed, via on-screen instructions, that they would hear pairs of voices in two blocks, by trait, and would be asked to make a decision regarding each pair. Participants were told that there was no time limit to their decision but were encouraged to answer with their first impression. After each pair of voices the question "Which voice did you perceive as more {dominant} {trustworthy}?" was displayed on the screen. Pressing the "s" key would mean that they perceived Voice 1 as being most dominant or trustworthy, whereas the "k" key represented Voice 2. The definitions of dominance and trustworthiness used in the instructions were "Dominance means having power and influence over others" and "Trustworthiness means able to be relied on as honest or truthful."

The order of the dominance and trustworthiness blocks, as well as the order in which the voice trials were presented within the block, were counterbalanced across participants. Male and female trials were presented randomly within the same block, as opposed to being presented in different blocks, to avoid an additional potential block-order effect caused by the gender of the voice. Finally, the order of the high- and low-pitched versions of the recordings within each trial was counterbalanced by including two trials of each pair in a block, with the high- and low-pitched versions in a different order. Therefore, within each block, the 20 pairs of voice samples were presented twice. The voices within each pair were played consecutively with a 1-s pause between the first voice and the second voice, and participants proceeded to the next trial by pressing “space.” The experiment lasted approximately 14 min.

Data analysis. Individual participants were used as the unit of analysis. To perform the analyses, each participant’s choices were coded as 1 if the low-pitched voice was selected and as 0 if the high-pitched voice was selected. The average of each participant’s choices, separately for each block, represents the proportion of trials in which the low-pitched voice was chosen over the high-pitched voice as more dominant or trustworthy. One sample *t*-tests were used to compare the proportion of trials in which the low-pitched voices were chosen as more dominant or trustworthy with a chance level of 0.5, which represents no preference for either low- or high-pitched voices. In cases where the data was not normally distributed, one-sample Wilcoxon signed-rank tests were used for the same purpose. All analyses were considered as two-tailed and tested at $\alpha = .05$.

Results

Dominance. The results showed that when judging vocal dominance, the low-pitched voices were selected more often compared with the high-pitched voices, both for male and for female voices (Figure 1). Participants selected the female low-pitched voices in 55% of the trials ($SD = 0.21$), and the male low-pitched voices in 62% of the trials ($SD = 0.21$), on average. A one-sample Wilcoxon signed-rank test (hypothesized median level = 0.5) revealed that the low-pitched male voices were chosen significantly more often than the chance level ($Z = 3.04$, $p < .002$). However, the preference for female low-pitched voices was not significantly higher than what would be expected by chance, $t(39) = 1.37$, $p < .178$.

Trustworthiness. Similar to dominance, when judging vocal trustworthiness, the low-pitched voices were selected more often compared with the high-pitched voices, both for male and for female voices (Figure 1). Participants selected the female low-pitched voices in 61% of the trials ($SD = 0.17$) and the male low-pitched voices in 59% of the trials ($SD = 0.22$), on average. A one-sample *t*-test (chance level = 0.5) revealed that the preference for low-pitched male voices was significantly higher than chance ($t(39) = 2.57$, $p < .014$). Similarly, a one-sample Wilcoxon signed-rank test showed that the preference for low-pitched female voices was also significantly higher than chance ($Z = 3.55$, $p < .001$).

Analysis by voice gender. A two-way ANOVA was used to explore a possible interaction between the trait examined and voice gender. A significant interaction between the two variables was found, $F(1,39) = 4.53$, $p < .04$, $partial-eta = 0.104$, which was driven by the preference for low-pitched voices in the dominance trials being stronger in response to male speakers (62%) compared with female speakers (55%), but the preference for low-pitched voices in the trustworthiness trials being stronger in response to female speakers

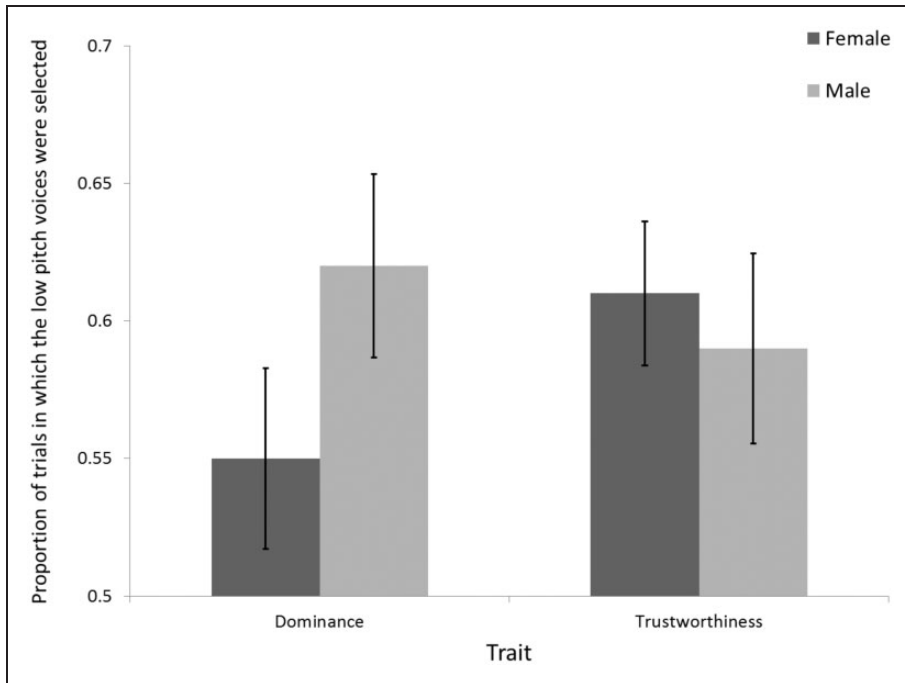


Figure 1. The proportion of female voice trials (dark gray) and male voice trials (light gray) in which the low-pitch version was selected, for both the dominance (left) and trustworthiness (right) traits. A mean value of 1 would indicate a 100% preference for low-pitched voices; 0.5 would indicate no preference for either low- or high-pitched voices; 0 would indicate 100% preference for high-pitched voices. Vertical axis is truncated between 0.5 and 0.7 for clarity. Error bars show standard error.

(61%) compared with male speakers (59%; Figure 1). Further analysis indicated that the difference between genders in the dominance trials was significant, $t(9) = 3.86$, $p < .004$, but was not significant in the trustworthiness trials. Additional analysis considering participant gender can be found in the Supplementary Information.

Analysis by order preference. To examine whether there was a tendency to select the first or the second voice, regardless of vocal pitch level, the proportion of trials in which participants selected the first voice over the second voice was calculated (Figure 2). In the dominance block, the first voice was selected on 54% of the female voice trials ($SD = 0.07$) and on 51% of the male voice trials ($SD = 0.14$), on average. In the trustworthiness block, the first voice was selected on 53% of the female voice trials ($SD = 0.15$) and on 51% of the male voice trials ($SD = 0.11$), on average. One-sample t -tests (chance level = 0.5) revealed that the first voice was selected significantly more often than chance only when judging the female speakers in the dominance block, $t(9) = 5.24$, $p < .001$.

Analysis by block order. To test for block-order effects, the proportion of trials in which the low-pitched voices were selected over the high-pitched voices, for the dominance and trustworthiness traits, was calculated separately both for participants who completed the dominance block first ($N = 20$), and for participants who completed the trustworthiness block first ($N = 20$). The preference for low-pitched voices in the dominance block was

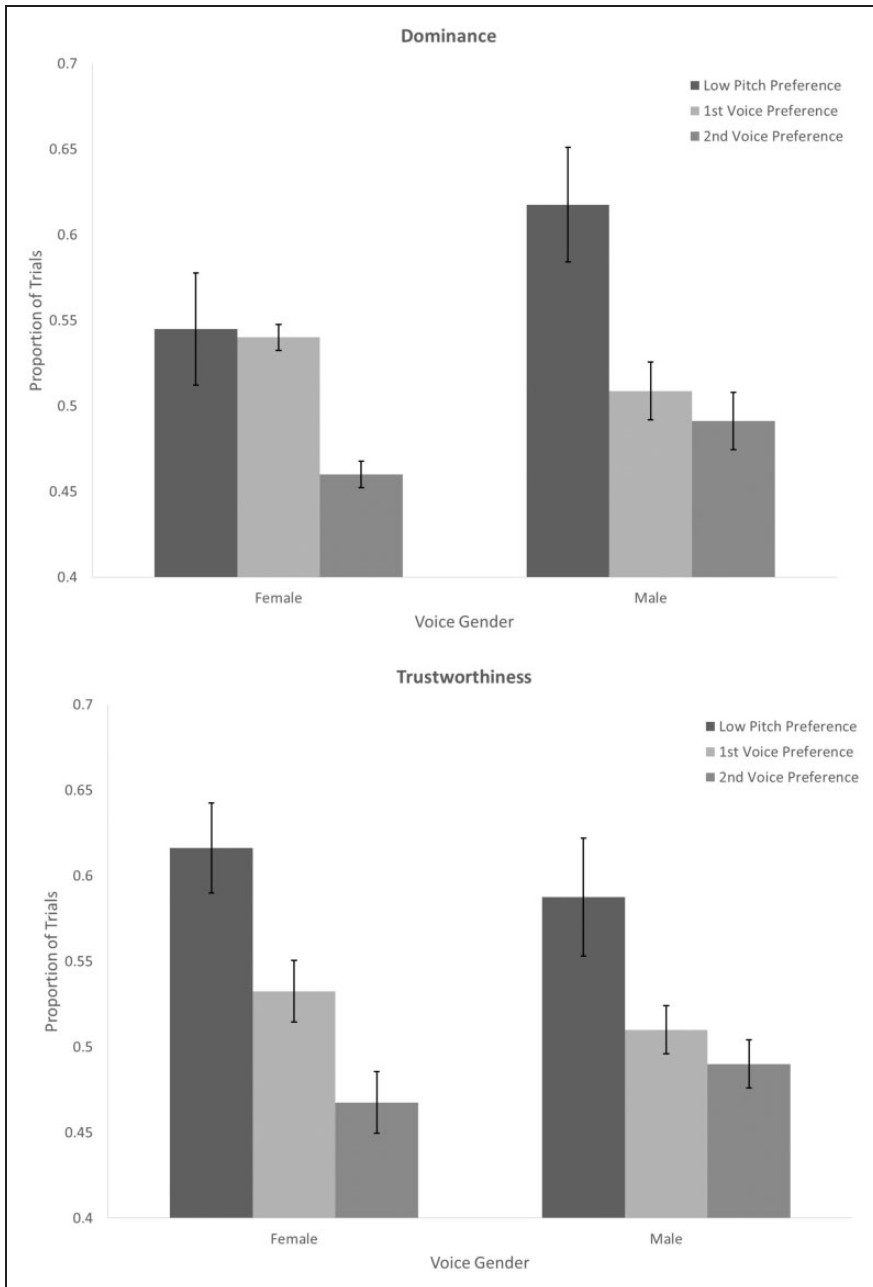


Figure 2. The proportion of female (left) and male (right) voice trials in which the low-pitch voice (dark gray) was selected, and the proportion of female and male voice trials in which the first (light gray) and second voice (mid gray) was selected, for the dominance (top graph) and trustworthiness (bottom graph) traits. A mean value of 1 would indicate a 100% preference for low-pitched voices; 0.5 would indicate no preference for either low- or high-pitched voices; 0 would indicate 100% preference for high-pitched voices. Vertical axis is truncated between 0.4 and 0.7 for clarity. Error bars show standard error.

60% for participants who completed that block first ($SD=0.17$) and 57% for participants who completed that block second ($SD=0.21$). The preference for low-pitched voices in the trustworthiness block was 61% when that block was completed first ($SD=0.18$) and 60% when that block was completed second ($SD=0.17$). Independent t -tests showed that these differences were nonsignificant (all $p > .05$).

Discussion

The results revealed a tendency to select the low-pitched voices as more dominant and more trustworthy compared with the high-pitched voices. The preference for low-pitched voices was significant for voices of both genders when judging vocal trustworthiness but for male voices only when judging vocal dominance. The strength of the preference was characterized by an interaction between the trait being judged and the gender of the voice; the preference was significantly stronger for male voices compared with female voices, in regard to dominance, but only marginally stronger for female voices compared with male voices, in regard to trustworthiness.

The finding that low-pitched male voices were judged as more dominant than high-pitched male voices is consistent with previous literature (Feinberg et al., 2006; Jones et al., 2010; McAleer et al., 2014; Puts et al., 2006; Puts et al., 2007; Tigue et al., 2012; Vukovic et al., 2011; Watkins et al., 2010; Wolff & Puts, 2010). One explanation for the link between low male vocal pitch and dominance is evident in the theoretical proposition that low pitch is misattributed to large body size and physical strength (Feinberg et al., 2005; Fitch, 1997). These characteristics may signal aggressive potential and have been previously correlated to physical dominance (Bartholomew & Collias, 1962; Puts et al., 2007; Puts et al., 2012). Furthermore, there is a perceived connection between low-pitched voices and assertiveness, charisma, and leadership skills, which are thought to be characteristics of social dominance (Ohala, 1983; Puts et al., 2007).

In regard to female voices, although low-pitched voices were more often judged as being more dominant compared with high-pitched voices, the results showed that this preference was not different from chance—or put simply, that there was no preference. This finding is in contrast to Borkowska and Pawlowski (2011) and Jones et al. (2010), who did find an overall significant preference for low-pitched voices using a similar task, and conversely in contrast to McAleer et al. (2014), who found that dominance was more associated with high vocal pitch in females. It has been proposed by Borkowska and Pawlowski (2011) that perceived female dominance is more closely related to social dominance, that is, characteristics such as assertiveness, charisma, and leadership skills (Bolinger, 1964; Ohala, 1983; Puts et al., 2007), than physical dominance, that is, aggressive potential (Bartholomew & Collias, 1962; Puts et al., 2007; Puts et al., 2012). It is notable that Puts et al. (2007) found that the manipulation of pitch in male voices had a stronger influence on the judgment of physical dominance than it did on the judgment of social dominance. It is possible that female dominance is more closely associated with social than physical dominance which could result in the influence of vocal pitch on judgments of dominance being weaker for female voices compared with male voices. This suggestion is in line with the finding by McAleer et al. (2014) that the variance in the perception of dominance from female voices is not well explained by vocal pitch, but is, in fact, better accounted for by the movement of the pitch in regard to glide and intonation. Our result of no preference can be elucidated by the preference toward selecting the first voice out of the female voice pairs when judging dominance, suggesting that responses were not being driven exclusively by vocal pitch.

Looking at trustworthiness, to our knowledge, the present study is the first to use a 2AFC task to explore the influence of vocal pitch level on judgments of trustworthiness where the stimuli are proposed as noncontextual but socially relevant speech. In contrast to our hypothesis, low-pitched male voices were perceived as more trustworthy compared with high-pitched voices. This supports Tigue et al. (2012), who used contextual stimuli, but is in contrast to McAleer et al. (2014), who used the same stimuli as the current study and found that higher-pitched male voices were rated as more trustworthy. Similarly, for female voices, in line with Klofstad et al. (2012), who used the 2AFC, the results of the present study showed that low-pitched female voices were perceived as more trustworthy compared with high-pitched female voices; yet, McAleer et al. (2014) found that intonation and not vocal pitch explained the variance of trustworthiness in female voices. To summarize, the current findings on trustworthiness are most consistent with previous studies that have used a different context but a similar task, as opposed to studies that have used a different task but a similar context.

As suggested, one possible explanation for the differing findings across studies is that the type of task used in the different experiments influences the relationship between vocal pitch level and trustworthiness judgments. McAleer et al. (2014) used a Likert task in which the degree of trustworthiness of the voices was rated individually, whereas the present study and others used a 2AFC where preference is judged between pairs of stimuli. It is possible that the preference for low-pitched voices reflects a general bias toward lower pitch when the task specifically requires comparison between two samples of the same speaker. Previous research has shown that additional stimuli in the immediate environment can alter percepts of personality traits for a target stimulus (Little, Burriss, Jones, DeBruine, & Caldwell, 2008; Re, Lefevre, DeBruine, Jones, & Perrett, 2014). This could, potentially, explain differences in studies using tasks with diverse cognitive demands. However, although studies using the 2AFC task frequently find preferences for low-pitched voices, there have been exceptions to this tendency (Apicella & Feinberg, 2009; Feinberg, DeBruine, Jones, & Perrett, 2008; Tigue et al., 2012). For example, specifically for male trustworthiness, Klofstad et al. (2012) found no significant influence of vocal pitch on trustworthiness judgments; yet, the current study showed that low-pitched male voices were perceived as more trustworthy. Both studies used a similar 2AFC task, which would reduce the possibility of a task effect.

An alternative explanation is that the relationship between vocal pitch and trustworthiness is nonlinear. Specifically, it is possible that high-pitched male voices are perceived as more trustworthy up to a certain pitch level above which they start to appear untrustworthy. A number of studies have revealed a link between deception and high vocal pitch (Apple, Streeter, & Krauss, 1979; Ekman, Friesen, & Scherer, 1976; Ekman, O'Sullivan, Friesen, & Scherer, 1991; Sporer & Schwandt, 2006; Lakhani & Taylor, 2003; Streeter, Krauss, Geller, Olson, & Apple, 1977; Taylor & Hick, 2007; Villar, Arciuli, & Paterson, 2013; Zuckerman, Koestner, & Driver, 1981). Borkowska and Pawlowski (2011) have reported a similar, nonlinear, relationship between vocal pitch and attractiveness in female voices. It is unclear what that optimal level of frequency for trustworthiness would be and, as such, has not yet been quantified in the literature. It is possible that the manipulated high-pitched male voices used in this study were higher than an optimal high-pitch level, but this would seem unlikely given the range of high and low pitch values for the stimuli reported in Table S1; there are examples where the low-pitched voice of one speaker is the preferred voice in that pairing but has a higher frequency than the unpreferred high-pitched voices of different voice pairings. Whether the relationship between trust and pitch is in fact nonlinear remains unclear.

A third alternative would be that the preference for low- and high-pitched voices is modulated by context and task, and that this effect is more pronounced in female voices than it is in male voices; for male voices an overall preference for low pitch may circumvent task and context. It would appear from the current study, and is supported by previous studies, that the preference for the low-pitched voice is consistent for male voices, with the majority of studies suggesting low pitch is perceived as dominant and to some degree trustworthy. For female voices, the effect is not as strong and the results are far more varied. Thus, it may be the case that people have an overall preference for low-pitched male voices, regardless of task and context, whereas the preference in female voices is more task and context-dependent.

Experiment 2

We performed a second experiment in order to explore the idea of an overall preference for low-pitched voices using the 2AFC task (High Pitch vs. Low Pitch), as in Experiment 1, with the additional variable of direction of speech (Forward vs. Backward) to address the effect of context at a basic level, that is, hearing a word that you clearly understand compared with hearing a word that you do not understand. If such a general bias exists, we would expect to find a preference for low-pitched voices irrespective of whether the stimuli are heard forward or backward. Based on previous findings, we hypothesized that a preference for low-pitched male voices regardless of speech direction would be found. For female voices, if context plays an important role as suggested by the results of Experiment 1, we would not expect to find a clear indication of an overall preference.

Methods

Participants. A total of 240 new participants (183 female) from the same participant pool as Experiment 1 took part in this experiment (average age = 20 years, $SD = 3$ years). No monetary incentive was given for taking part; participants were awarded partial course credit where appropriate.

Stimuli and apparatus. For this experiment, the 10 male and 10 female voice pairings (High Pitch vs. Low Pitch) from the dominance version of Experiment 1 were used, to remove any potential issue with the repetition of sounds across conditions. After piloting, one female voice was replaced due to a strong potential for familiarity by the listeners due to the speaker's role within the local community. Stimuli were normalized for loudness from the original recordings after the replacement female sound was included. It is worth noting that the pitch values would have been similar had we selected the stimuli used for the trustworthiness judgments in Experiment 1, as voices were selected from the mid ranges of the two-dimensional space where dominance and trustworthiness intersect.

Procedure. The experiment took place in the experimental laboratories at the University of Glasgow. Participants were required to complete a 2AFC task during which they listened to pairs of voices comprising high- and low-pitched versions of the original recordings. General procedures and timings were the same as in Experiment 1 except that after each pair of voices was heard, the question "Which voice do you prefer?" was displayed on the screen. Pressing the "1" key would mean that they preferred the first voice, whereas "9" would indicate that they preferred the second voice. No other information was given to the participants about the context of the voices other than that they had to choose one of the voices out of each pair.

Participants were split into four counterbalancing conditions of the two independent variables (direction of speech and gender of voice), run across a 1-week time period. Direction of speech was run as a between-subjects variable and gender of voice was run as a within-subjects variable. Male and female voices were presented in separate gender blocks and the order of the blocks was counterbalanced across participants. This resulted in four conditions: backwards speech with male voices first; backwards speech with female voices first; forwards speech with male voices first; and forwards speech with female voices first.

Finally, as in Experiment 1, trials were presented randomly within each block, and the order of the high- and low-pitched versions of the recordings within each trial was counterbalanced by including two trials of each pair in a block, with the high- and low-pitched versions in a different order. Therefore, within each block, the 20 pairs of voice samples were presented twice (40 trials). The voices within each pair were played consecutively with a 1-s pause between the first voice and the second voice. The experiment moved on automatically to the next trial after a response was given, with a one second gap between voice pairings. The experiment lasted approximately 14 min.

Results

Analysis by block-order. An initial three-way ANOVA (gender of voice: Male vs. Female; direction of speech: Forward vs. Backward; gender heard first: Male vs. Female) was conducted to test for possible order effects attributed to the gender of the voices in the block that was presented first to participants. No significant interactions or main effect involving gender heard first were found, and therefore all results were collapsed across this variable.

Analysis of overall preference. The overall preference for low-pitched voices in male and female speakers using forward and backward speech was analyzed to determine whether there is a general low pitch preference in both genders regardless of direction of speech. As shown in Figure 3, the preference for low-pitched voices appears stronger in male voices than in female voices, but does not appear to be influenced by direction of speech. This was confirmed with a two-way ANOVA comparing gender of voice (Male vs. Female) and direction of speech (Forward vs. Backward) which showed a main effect of gender of voice only, $F(1,238)=31.8$, $p < .01$, $partial-eta=0.12$. Neither the main effect of direction of speech nor the interaction between the two variables was found to be significant (all F 's < 1)

Analysis by voice gender. We next analyzed the four conditions separately to see if the preference for low-pitched voices was significantly different from the chance level (0.5 = no preference). For the female voices, the preference for low pitch in the Forwards Speech ($M=0.52$, $SD=0.16$, confidence interval [CI]=0.49 to 0.55) and in the Backwards Speech ($M=0.53$, $SD=0.20$, CI=0.49 to 0.57) conditions was not significantly different from the chance level of 0.5; despite showing mean preference values indicating a low pitch preference, both conditions were only marginally higher than 0.5, and the 95% confidence intervals crossed the chance level boundary. In contrast, for the male voices. the low pitch preference in the Backward Speech condition ($M=0.61$, $SD=0.21$, CI=0.57 to 0.65) was found to be significantly different from the chance level of 0.5, $t(119)=5.56$, $p < .01$. Lastly, data in the Forward Speech condition with male voices ($M=0.60$, $SD=0.22$) on analysis failed a Kolmogorov–Smirnov test for normality, $p=.004$, and thus a one-sample Wilcoxon signed-rank test was performed (hypothesized median=0.5); the preference for low pitch was found to be significantly different from chance ($Z=4.667$, $p < .01$, median=0.61,

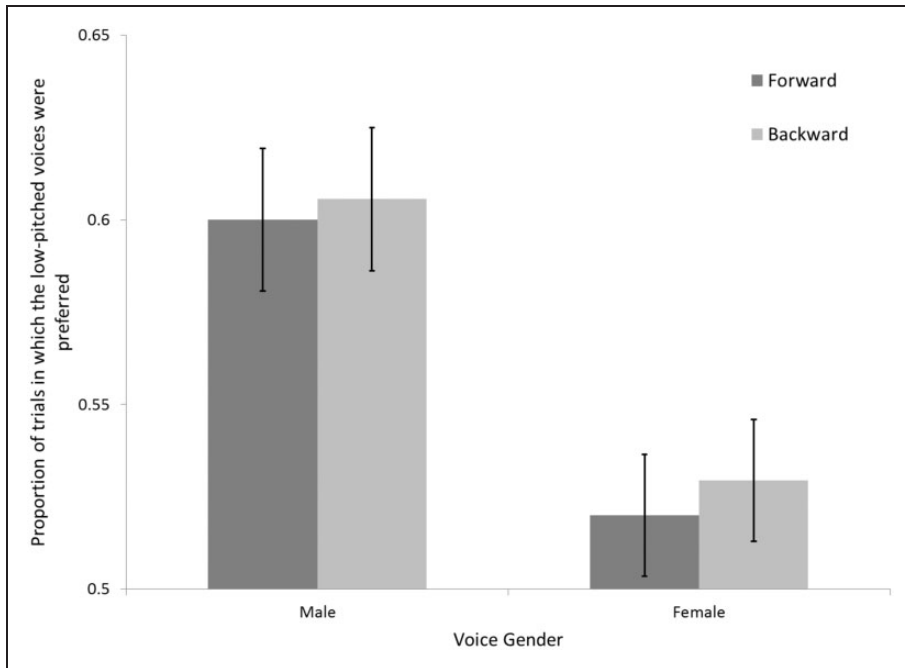


Figure 3. The proportion of male voice trials (left) and female voice trials (right) in the forward condition (dark gray) and backward condition (light gray) in which the low-pitch version was selected. A mean value of 1 would indicate a 100% preference for low-pitched voices; 0.5 would indicate no preference for either low- or high-pitched voices; 0 would indicate 100% preference for high-pitched voices. Vertical axis is truncated between 0.5 and 0.65 for clarity. Error bars show standard error.

Bootstrap CI = 0.55 to 0.69). Additional analysis considering participant gender can be found in the Supplementary Information.

Discussion

Experiment 2 was performed to explore the question of whether there was an overall preference for low-pitched voices in male and female speakers, and whether this was affected by speech content. Evidence was found to suggest that, at least for male voices, there is an underlying preference for the low-pitched voice, and this is irrespective of what is said and the gender of the listener. The effects appear to be small but consistent. In contrast, for female voices, there is no strong evidence to suggest a basic underlying preference for low pitch. All means for the female contrasts were in favor of a preference for low-pitched voices; however, the results were not significantly different from chance, suggesting a larger range of preference across listeners. In turn, we found no evidence that this range was modulated by the gender of the listener (see Supplementary Information) or of the utterance heard.

Across the two experiments we found a consistent preference for low-pitched male voices either when people were asked to determine which is most trustworthy or most dominant, or, simply, which is preferred. Again, this appears consistent irrespective of the gender of the listener, and of what was heard. The finding for male voices is consistent with previous studies (e.g., Jones et al., 2010; Klofstad et al., 2012; Tigue et al., 2012), and promotes the understanding that this preference for low-pitched male voices transcends experimental

scenarios and vocalizations. Why such a default exists is, potentially, related to relationships between perceived health, reproductive success, strength, and decreasing voice pitch (Apicella et al., 2007; Oosterhof & Todorov, 2008), or to a proposed “soothing” or comforting nature of low-pitched voices; a relationship shown to exist even in animals (McConnell, 1990; Heleski et al., 2015). However, the results, at least for judgments of trustworthiness, are in contrast with McAleer et al. (2014), who, using a larger set of the same stimuli and a Likert scale for each individual voice, showed a higher pitch to be more trustworthy. Additional stimuli in the environment have previously been found to influence the signals used to make a judgment of trustworthiness from voices (Jones et al., 2008; Vukovic et al., 2008). This suggests that the specific task and stimuli set can create a degree of bias, however small, toward male low-pitched voices, and studies using a 2AFC should consider that their chance level is naturally shifted toward the low-pitched voice. To fully understand this claim, however, a direct comparison of tasks is the next step for future research.

For females, across both experiments, the preference for low pitch is only evident when specifically judging which voice is most trustworthy, and not found for dominance judgments or general preference decisions. Klofstad et al. (2012) also found a preference for low-pitched female voices when judging trustworthiness in a political scenario, and these studies combined would indicate a consistency in this relationship irrespective of scenario or utterance as long as the decision relates to trustworthiness. It is possible that this connection is driven by a relationship between age and trust, with older voices tending to have a lower pitch, at least in middle age, and older people being perceived as more trustworthy (Klofstad, Anderson, & Nowicki, 2015). A lack of consistency between this finding and McAleer et al. (2014), who did not find a strong relationship between pitch and trustworthiness in females, is perhaps not that surprising considering that in the present study the key acoustical measures, that is, glide and intonation, would have been constant across pairs of voices in each trial, and thus, the judgment could only be made on pitch.

In regard to dominance in female voices, the finding of no pitch preference may indicate that such judgments concerning female voices are partly modulated by context, or at least more so by context, when compared with male voices. This explanation is supported by the finding of no overall pitch preference for female voices in Experiment 2, suggesting that pitch preferences in female voices are more driven by context or situation than by a general overall bias toward low pitch. Furthermore, the tendency to respond to the first voice in each pair, when judging dominance in Experiment 1, would again indicate a lack of an underlying cognitive or perceptual bias driving the participants’ decisions. However, we would propose that it is not the case that there is never a preference for the low pitch in female voices, but that the preference for whether people prefer high- or low-pitched female voices is contextually or situationally dependent (e.g., trustworthiness in low-pitched female voices in Experiment 1 and Klofstad et al., 2012). Overall, we suggest that a lack of general bias may help explain inconsistencies found when comparing previous studies that have attempted to relate high or low pitch to dominance in female voices (Borkowska and Pawlowski, 2011; Jones et al., 2010; Ohala, 1982). Moreover, we propose that pitch preference in female voices is driven by a combination of situation and listener as yet to be fully understood.

One potential caveat in our findings would be that in terms of presentation of stimuli, Experiment 1 randomized male and female pairings within blocks, whilst Experiment 2 separated blocks by gender. Although it is possible that this may have influenced results, we believe that similarities between findings of Experiment 1 and previous studies that have blocked trials by gender (e.g., Klofstad et al., 2012) would suggest that in the 2AFC task focus is directed to the paired voices being considered in that trial and not in consideration of

all voices in that block; as is a purpose of the 2AFC task as opposed to the Likert scale used in other studies (e.g., McAleer et al., 2014). That said, given that it is known that additional stimuli in the general environment could potentially have an influence (Little et al., 2008; Re et al., 2014), further analysis comparing the different tasks and stimuli environment should be considered.

General Conclusions

The initial aim of the present study was to explore the influence of vocal pitch level on first-impression judgments of dominance and trustworthiness across genders using noncontextual stimuli. When judging dominance, the preference for low-pitched male voices was much stronger compared with the preference for low-pitched female voices, which failed to reach significance. When judging trustworthiness, the strength of the preference for low-pitched voices as more trustworthy was on a similar level for both genders. The removal of context through reversed speech suggested a general preference for low-pitched voices; stronger in males than in females. This claim was tested in the second experiment of the study and revealed that in male voices there does indeed appear to be an overall preference for low pitch regardless of context or personality trait. For female voices, no overall preference for low or high pitch was found, suggesting that whether people prefer a higher- or lower-pitched female voice is dependent on the context and situation. However, the possibility that different findings may arise by changing the task or contrasting voices of different speakers, as opposed to a high and low version of the same speaker (as is the standard paradigm in this study and related others), cannot be ruled out on the basis of the findings of this study.

In conclusion, it is proposed that the perception of low-pitched voices as more dominant and more trustworthy, when making first-impression judgments, is partly driven by a default bias. This bias appears stronger for male voices than for female voices, in which case context, or an as yet untested variable, influences the judgment. Overall, we believe that the findings of the present study: reiterate the fundamental role of vocal pitch in the formation of first impressions from voice; give important ground truth knowledge for future research in this field; and, in turn, may influence future voice-driven systems with the suggestion to fully consider how the voice utilized in their technology will be heard and judged.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proceedings of the Royal Society B: Biological Sciences*, 276, 1077–1082.
- Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, 3, 682–684.
- Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37, 715–727.

- Bartholomew, G. A., & Collias, N. E. (1962). The role of vocalization in the social behaviour of the northern elephant seal. *Animal Behaviour*, *10*, 7–14.
- Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, *102*, 711–725.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, *5*, 341–345.
- Bolinger, D. (1964). Intonation as a universal. In *Proceedings of the 5th Congress of Phonetics, Cambridge 1962* (pp. 833–848).
- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*, 55–59.
- Dabbs, J. M. Jr., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*, 801–804.
- Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica*, *16*, 23–27.
- Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. R. (1991). Invited article: Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior*, *15*, 125–135.
- Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*, 783–788.
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception*, *37*, 615–623.
- Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., . . . Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, *49*, 215–222.
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*, 561–568.
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, *102*, 1213–1222.
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*, 258–267.
- Fleming, D., Giordano, B., Caldara, R., & Belin, P. (2014). Language-familiarity effect for speaker discrimination without comprehension. *Proceedings of the National Academy of Sciences*, *111*, 13795–13798.
- Ferdenzi, C., Patel, S., Mehu-Blantar, I., Khidasheli, M., Sander, D., & Delplanque, S. (2013). Voice attractiveness: Influence of stimulus duration and type. *Behavior Research Methods*, *45*, 405–413.
- Heleski, C., Wickens, C., Minero, M., DallaCosta, E., Wu, C., Czeszak, E., & Köenig von Borstel, U. (2015). Do soothing vocal cues enhance horses' ability to learn a frightening task? *Journal of Veterinary Behavior: Clinical Applications and Research*, *10*, 41–47.
- Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A., & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *The Journal of the Acoustical Society of America*, *106*, 1532–1542.
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2008). Integrating cues of social interest and voice pitch in men's preferences for women's voices. *Biology Letters*, *4*, 192–194.
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*, 57–62.
- 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 B-Biological Sciences*, *279*, 2698–2704.
- Klofstad, 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*, e0133779.

- Lakhani, M., & Taylor, R. (2003). Beliefs about the cues to deception in high- and low-stake situations. *Psychology, Crime & Law*, *9*, 357–368.
- Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*, *21*, R143–R145. doi:10.1016/j.cub.2010.12.033
- Little, A. C., Burriss, R. P., Jones, B. C., DeBruine, L. M., & Caldwell, C. A. (2008). Social influence in human face preference: men and women are influenced more for long-term than short-term attractiveness decisions. *Evolution and Human Behaviour*, *27*, 140–146.
- McAleer, P., Todorov, A., & Belin, P. (2014). How do you say ‘Hello’? Personality impressions from brief novel voices. *PLoS One*, *9*, e90779doi:10.1371/journal.pone.0090779
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social-perception. *Psychological Review*, *90*, 215–238.
- McConnell, P. B. (1990). Acoustic structure and receiver response in domestic dogs, *Canis familiaris*. *Animal Behaviour*, *39*, 897–904.
- Nass, C., & Lee, K. (2001). Does computer-synthesised speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, *7*, 171–181.
- Newman, S. R., Butler, J., Hammond, E. H., & Gray, S. D. (2000). Preliminary report on hormone receptors in the human vocal fold. *Journal of Voice*, *14*, 72–81. doi:10.1016/S0892-1997(00)80096-X
- Ohala, J. J. (1982). The voice of dominance. *The Journal of the Acoustical Society of America*, *72*, S66–S66.
- Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, *40*, 1–18.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*, 11087–11092.
- Puts, D. A., Apicella, C. L., & Cardenas, R. A. (2011). Masculine voices signal men’s threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 601–609.
- Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*, 283–296.
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men’s voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, *28*, 340–344.
- Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual Selection on Human Faces and Voices. *Journal of Sex Research*, *49*, 227–243.
- Re, D. E., Lefevre, C. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2014). Impressions of dominance are made relative to others in the visual environment. *Evolutionary Psychology*, *12*, 251–263.
- Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. P. L. Tagiuri (Ed.), *Person perception and interpersonal behaviour* (pp. 300–315). Palo Alto, CA: Stanford University Press.
- Scherer, K. R. (1972). Judging personality from voice—cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality*, *40*, 191–210.
- Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*, 421–446.
- Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. *Journal of Personality and Social Psychology*, *35*, 345–350.
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*, 105–118.
- Taylor, R., & Hick, R. F. (2007). Believed cues to deception: Judgments in self-generated trivial and serious situations. *Legal and Criminological Psychology*, *12*, 321–331.
- 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*, 210–216.

- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America*, *85*, 1699–1707.
- Villar, G., Arciuli, J., & Paterson, H. (2013). Vocal pitch production during lying: Beliefs about deception matter. *Psychiatry, Psychology and Law*, *20*, 123–132.
- Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, A. C., & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences*, *45*, 451–456.
- Vukovic, J., Jones, B. C., Feinberg, D. R., DeBruine, L. M., Smith, F. G., Welling, L. L. M., & Little, A. C. (2011). Variation in perceptions of physical dominance and trustworthiness predicts individual differences in the effect of relationship context on women's preferences for masculine pitch in men's voices. *British Journal of Psychology*, *102*, 37–48.
- Watkins, C. D., Fraccaro, P. J., Smith, F. G., Vukovic, J., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2010). Taller men are less sensitive to cues of dominance in other men. *Behavioral Ecology*, *21*, 943–947.
- Wolff, S. E., & Puts, D. A. (2010). Vocal masculinity is a robust dominance signal in men. *Behavioral Ecology and Sociobiology*, *64*, 1673–1683.
- Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social & Personality Psychology Compass*, *2*, 1497.
- Zuckerman, M., & Driver, R. E. (1989). What sounds beautiful is good—The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, *13*, 67–82.
- Zuckerman, M., Koestner, R., & Driver, R. (1981). Beliefs about cues associated with deception. *Journal of Nonverbal Behavior*, *6*, 105–114.