Vocal attractiveness in homosexual and heterosexual listeners

Belyk, M.[1], Horak, L.[2] Beone, C.[2], and McGettigan, C.[2]

[1] Edge Hill University, Department of Psychology
[2] University College London, Department of Speech Hearing and Phonetic Sciences

[PREPRINT NOT YET PEER REVIEWED]

Corresponding author:

Michel Belyk, Ph.D.
Edge Hill University
St Helens Road
Ormskirk, UK
L39 4QP
belykm@edgehill.ac.uk

Abstract

The human voice is a sexually dimorphic trait that signals biologically and socially relevant information about the person who is speaking, including how attractive they sound to listeners. Biological factors that index reproductive fitness lead a voice to be perceived as attractive by heterosexual listeners. However, such a monolithic view may belie the role of social influences in a dynamic signalling system that is under extensive voluntary control. To the extent that homosexual listeners may not share the same vested interest in assessing reproductive fitness, they provide a valuable lens through which to dissociate biological from social influences on vocal attractiveness. In this study, equal numbers of female and male heterosexual and homosexual participants rated the attractiveness of voice recordings that had been masculinised or feminised, either synthetically by digital manipulation of recorded speech tokens (Experiment 1) or naturalistically by volitional modulation during speech (Experiment 2). We replicated previous findings that heterosexual listeners prefer voices that exaggerate rather than attenuate sexually dimorphic vocal cues. However, data from homosexual listeners provides a richer context, indicative of an overall bias for masculinised voices. Despite differences in how listeners responded to voice modulations, all groups displayed a global preference for unmodulated baseline voices that provide a more honest representation of the speaker. Likewise, there was a strong consensus across sexual orientations regarding which speakers had the most attractive unmodulated voices. Listeners maintain separate sets of preferences based on stable features of the voice that are driven by principles of biological signalling versus the dynamic social signals that it also carries. These findings demonstrate the added scientific value of widening the inclusivity of research.

Keywords
vocal attractiveness; voice modulation; vocal trait judgement; sexual orientation, biological signalling; social signalling; vocal tract length; vocal honesty

The voice is a common carrier for a wide range of communicative signals. The most obvious of these signals are speech, song, and the expression of emotion. However, a voice also carries socially relevant information about the person to whom it belongs, such as the attractiveness, dominance, and trustworthiness of a speaker (Jody Kreiman & Diana Sidtis, 2011; Lavan & McGettigan, 2023; Scott & McGettigan, 2016). Listeners make these judgements rapidly and with high agreement, even from very brief samples of speech, suggesting that they decode these signals using a shared library of acoustical cues (Lavan, 2023; McAleer et al., 2014; Mileva & Lavan, 2023). In the case of vocal attractiveness, at least some of those cues - namely, vocal pitch and apparent vocal size - have been identified as potential indicators of secondary sexual characteristics of the body to which the voice belongs.

Both vocal pitch and apparent vocal size are sexually dimorphic traits that index anatomical or physiological attributes of the speaker (Hughes et al., 2002). The pitch of the voice is the perceptual correlate of its fundamental frequency (f0), the frequency at which the vocal folds of the larynx vibrate to produce the sound of a given person's voice at any given time (Story & Titze, 1995; Titze, 2008). Vocal pitch scales inversely with vocal fold length and thickness, and is generally higher in children than in adults, and higher in adult females than adult males. The percept of vocal size relates to the overall size of the vocal tract, which forms a resonant cavity that selectively amplifies or dampens certain frequency bands within the voice. Longer vocal tracts selectively amplify the lower frequency harmonics of the voice, thus the voices of adults sound larger than children's voices, and adult male voices sound larger than adult female voices.

As secondary sexual characteristics, vocal cues signal biologically relevant information about the individual such as their current and historical hormone profiles. They form part of a multimodal display of mate quality (Feinberg, 2008; Groyecka et al., 2017) in which the voice contributes information that is either complementary (Feinberg et al., 2008; Fraccaro et al., 2010; Lander, 2008; Saxton et al., 2006; Valentova et al., 2017) or redundant (Feinberg, 2008; Groyecka et al., 2017) with facial attractiveness. Indeed the tissue of the vocal folds contain receptors for the sex hormones testosterone, estrogen, and progesterone (Brunings et al., 2013; Newman et al., 2000; Voelter et al., 2008), and correspondingly hormone related changes in the larynx are detectable in the spoken voice. For example, changes in the size of the vocal folds as well as the descent of the larynx at puberty distinguishes the voices of children from adults, with a secondary descent of the larynx in males creating further sexual dimorphism (Fitch & Giedd, 1999; Kahane, 1978). In females, hormonal changes across the menstrual cycle, menopause, and during pregnancy may have more dynamic effects on voice across the lifespan (Abitbol et al., 1999; Zamponi et al., 2021). The physiological consequences of such hormone variation – between and within individuals - are made audible to listeners by their effects on f0 and apparent vocal size (Dabbs & Mallinger, 1999; Evans et al., 2008; Harries et al., 1997; Markova et al., 2016). These hormonal changes have a corresponding influence on the perceived attractiveness of a voice (Babel et al., 2014; McAleer et al., 2014; Pisanski et al., 2014; Xu et al., 2013).

Playback experiments typically report that voices are perceived to be more attractive when they exaggerate sexually dimorphic acoustical features. Female voices that are more feminine (i.e., high pitched, small apparent vocal size) are more attractive to heterosexual men (Borkowska & Pawlowski, 2011; Feinberg et al., 2005; Fraccaro et al., 2010), while male voices that are more masculine (i.e., low pitched, large apparent vocal size) are more attractive to heterosexual women (Apicella & Feinberg, 2009; Feinberg et al., 2005; Feinberg, 2008; Pisanski & Rendall, 2011; Šebesta et al., 2017).

However, previous research has been based disproportionately on heterosexual courtship – the perceptions of attractiveness in opposite-sex voices as potential mates, versus same-sex voices as competitors. The exclusion of homosexual listeners has typically been justified either on the grounds that 1) vocal attractiveness research is concerned primarily with how these systems of communication promote the reproductive fitness of both the signaller and the receiver, which are of less direct relevance to homosexual courtship, or 2) vocal signalling and receiving may manifest differently in this population such that samples inclusive of homosexual participants may obfuscate patterns of behaviour that are only strongly predicted for heterosexual populations.

One previous study demonstrated that voice preferences of androphilic listeners for masculine male voices differed subtly between heterosexual females and homosexual males (Valentová et al., 2013). While both groups were more attracted to voices that were rated as sounding more masculine, this effect was driven predominantly by single homosexual men more than coupled men, and by coupled heterosexual women more than single women. The similarity in the direction of preference suggests that voice preferences may be at least partially determined by the biological sex of the speaker, though the influence of social factors is also evident. This is supported by parallel research in facial attractiveness demonstrating, for example, that while heterosexual males have the predicted preference for feminised female faces, homosexual men prefer masculinised male faces and homosexual women prefer masculinised female faces (Glassenberg et al., 2010; Shiramizu et al., 2020). The complexity of these findings further demonstrates the interplay of biological and social factors.

Nonetheless, the voice is a highly dynamic social signal, and individual speakers can produce pitches spanning several octaves, well beyond the mean sex differences in vocal pitch. Likewise, apparent vocal size can be modulated, for example by raising or lowering the larynx in the airway to either elongate or shorten the vocal tract in simulation of a taller or shorter speaker (Belyk & Brown, 2017; Pisanski, et al., 2016). Speakers can utilise these mechanisms to volitionally modulate their apparent vocal size and influence how large they sound (Belyk et al., 2022; Pisanski & Reby, 2021). These behaviours influence how attractive speakers sound to listeners (Fraccaro et al., 2013; Hughes et al., 2014) and occur spontaneously in courtship contexts (Fraccaro et al., 2011; Pisanski et al., 2018). Notably, listeners do have some capacity to detect vocalisations that contain potentially deceptive vocal cues and attempt to compensate for them perceptually, even if incompletely (Pisanski & Reby, 2021).

In light of the potential contribution of both biological and social factors through the interacting influences of sex, sexual orientation, and voice modulation, on listener's voice preferences, we tested four hypotheses about vocal attractiveness perception:

1. The listener hypothesis: Female listeners will perceive low-pitched and large-sounding voices to be more attractive (vice versa for male listeners).

2. The speaker hypothesis: Listeners who are attracted to females will perceive high-pitched and small-sounding voices to be more attractive (vice versa for listeners who are attracted to males).
3. The variability hypothesis: There will be greater individual variation among homosexual listeners, indicative of the greater influence of social factors in this population.
4. The naturalism hypothesis: Listeners will perceive modulated voices to be less attractive than unmodulated voices.

Speakers were selected from a corpus of recordings in which the speakers' voices were modulated in vocal pitch and vocal size along a continuum from masculinised to feminised (Waters, Kanber et al., 2021). Female and male listeners who identified themselves as heterosexual or homosexual rated the attractiveness of voices of the sex to which they were attracted. In Experiment 1 listeners rated voices that were modulated by digital synthesis, while in Experiment 2 they rated voices that were modulated volitionally by the speakers.

Methods

The design, procedure, hypotheses, and analyses for both experiment 1 (https://osf.io/w63au) and experiment 2 (https://osf.io/h8cg7) were preregistered. Any deviations from preregistered study plans are explained below.

Participants

Experiment 1

A total of 180 adults were recruited from Prolific. Participants were divided evenly between sexes and sexual orientations (45 heterosexual women, 45 homosexual women, 45 heterosexual men, and 45 homosexual men), aged between 18-40 years (M = 27.2 years, SD 5.9), with self-reported normal or corrected to normal vision, no hearing difficulties, English as a first language, and no ongoing mental health conditions. All recruited participants had high approval ratings within the recruitment platform (>90). Participants were recruited through separate recruitment drives restricted by sex and sexual orientation to reduce the influence of priming and selection bias. Other sexual orientations were not sampled in this study. The study was approved by the Research Ethics Committee at the Department of Speech, Hearing and Phonetic Sciences, University College London (SHaPS-2019-CM-030). Participants provided written informed consent and were reimbursed at a rate of £7.50 per hour, prorated for an estimated 15-minute completion time.

Experiment 2

A separate pool of 180 adults was recruited following the same demographic profile and recruitment procedures (M = 26.6 years, SD 6.3).

Stimuli

Overview

Stimuli were selected from among a set of recordings generated by previous speech production experiments (Belyk et al., 2022; Waters et al., 2021). Speakers were recorded producing the carrier words "bead" and "bard" 5 times each at a slow pace using their habitual voice. These words were chosen because they are framed by the same stop consonants, which facilitate the detection of sound onset and offset, and are distinguished by the front vowel /i/ or the back vowel /ɑ/ which are highly contrastive in English.

Representative tokens were selected for each speaker with consistent duration (0.6-0.8 s) durations with clear voice quality (i.e., without vocal fry which may induce distortions during subsequent signal processing). Tokens were then independently transformed in f0 and apparent Vocal Tract Length (aVTL) using procedures developed by Chris Darwin at the University of Sussex (http://www.lifesci.sussex.ac.uk/home/Chris_Darwin/Praatscripts/VTchange). A central, "baseline voice" version of each word was produced, in which the formants were unchanged but the f0 was shifted 2 semitones upward from the original. This adjusted baseline was needed to compensate for the greater vocal range that is available to most speakers in raising relative to lowering vocal pitch. Manipulated versions of each carrier word were further created by independently shifting vocal pitch and formant frequencies (which determines aVTL) by +/- 2, or +/- 4 semitones relative to baseline.

Speakers were provided with training in how to independently raise and lower vocal pitch and how to raise or lower their larynx within the vocal tract to exaggerate or attenuate their apparent vocal size. See https://osf.io/6pqkt/ for training materials and summaries of acoustical measurements, however audio recordings are not provided in order to preserve the anonymity of the speakers. Importantly, participants were given no instructions in regards to sounding more or less attractive, though vocal pitch and vocal size are strong drivers of vocal attractiveness. After training, speakers completed an imitation task in which they listened to the synthetically manipulated recordings of their own voice and imitated them attempting to match them as closely as possible. Hence the synthetic stimuli acted as guides to facilitate the speakers' attempts at exaggerating or diminishing their vocal size. On the basis of these experiments, recordings from the 5 females and 5 males who were most effective at vocal modulation (Good Vocal Modulators) and the 5 females and 5 males who were least effective at vocal modulation (Poor Vocal Modulators) were selected as stimuli for the current experiment (see Belyk et al. 2022 for further details on how vocal modulation ability was evaluated).

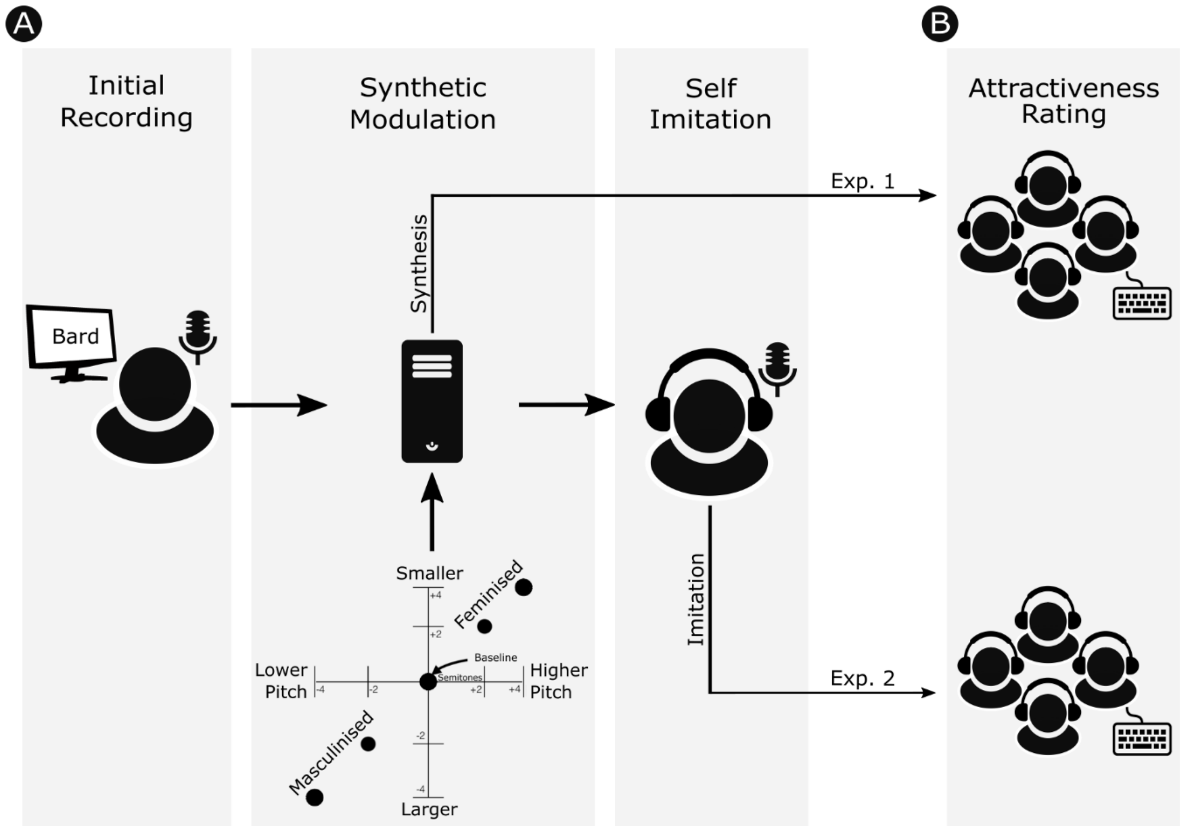

Figure 1: *Overview of methodology. A) Previous study (Waters, Kanber et al., 2021): Speakers were recorded as part of a separate study. An initial set of recordings from each speaker had f0 and formants acoustically manipulated along an axis from feminised to masculinised. Speakers then attempted to imitate these synthetic recordings in a sound attenuated booth. B) Current study: In Experiment 1 listeners rated the attractiveness of voices manipulated by synthesis, while in Experiment 2 listeners rated the attractiveness of voices modulated by imitation.*

Experiment 1

For each speaker, tokens of *synthetically manipulated* vocalisations spanning -4, -2, 0 ,+2, +4 semitones of both f0 and aVTL were selected as stimuli for experiment 1. These recordings span a continuum from masculinised voices (-4, -2 semitones) through baseline (0 semitones) to feminised voices (+2, +4 semitones).

Experiment 2

For each speaker, tokens of *imitated* vocalisations in which speakers attempted to match auditory targets spanning -4, -2, 0, +2, +4 semitones of both f0 and VTL were selected as stimuli for experiment 2. These recordings span a continuum from masculinised (-4, -2 semitones) through baseline (0 semitones) to feminised (+2, +4 semitones). These differ from Experiment 1 in that they were produced by living vocal tracts and subject to biological constraints not evident in the synthetic stimuli. By design, the 10 speakers spanned a range of abilities in this vocal matching task as assessed by Belyk et al. (2022).

Procedure

Experiment 1 and Experiment 2 followed identical procedures save differences in experimental stimuli. In Experiment 1 participants rated synthetically manipulated recordings whereas in Experiment 2 participants rated imitations.

Participants were recruited for online data collection via Prolific. At recruitment participants were instructed that they should complete the study on a computer while wearing headphones in a quiet environment. An initial headphone screening task ensured compliance with these instructions (Milne et al., 2020). Participants who did not pass the headphone screening were excluded from further data collection and a replacement was recruited. Participants were also made aware that the experiment would include 5 catch trials intended to monitor for low quality respondents. Catch trials consisted of a synthesised voice instructing participants to respond with an arbitrary number within the range of slider responses utilised in the main body of the experiment. Participants who produced errors on more than one trial were excluded from further analysis and a replacement was recruited.

Participants were presented with individual vocal stimuli in isolation and instructed to rate how attractive they found each voice using a slider ranging from 0 (not attractive) to 100 (very attractive). Above the response slider, the labels “moderately unattractive” and “moderately attractive” provided intermediate anchoring points for the scale. Each participant listened to stimuli from 10 speakers of the sex to which they were attracted using 2 carrier words (“Bead” or “Bard”) at 5 voice manipulation levels along a continuum from masculinised to feminised and centred on baseline (-4, -2, 0, +2, +4 Semitones) to a total of 10 (speakers) x 2 (words) x 5 (voice manipulations) = 100 trials presented in random order.

Analysis

Vocal attractiveness ratings were analysed by linear mixed models (Bates et al., 2015) using R (R Core Team, 2019). Models included fixed effects of 1) vocal manipulation condition, 2) the sexual orientation of the listener and 3) the sex of the speaker. Models included a random intercept of participant and a random intercept of speaker. This random effects structure deviates slightly from the preregistered form in which a planned random slope was nested within participant, however the planned model suffered from singular fit and it was therefore necessary to simplify the models (see equation 1).

Eq (1) *Attractiveness ~ Sexual_Orientation * Voice_Sex * Voice_Modulation + (1|Listener) + (1|Speaker)*

Test statistics are reported alongside standardised (Gelman, 2008; Gelman & Yo, 2015) effect estimates (E) and their confidence intervals (CI). These estimates model differences from a baseline which was arbitrarily selected as 0 Semitones of modulation in female speakers as rated by heterosexual (i.e., male) listeners. Hence estimates indicate change in vocal attractiveness ratings for modulated relative to unmodulated voices, male speakers relative to female speakers, and homosexual listeners relative to heterosexual listeners.

Results

Experiment 1: Modulation by synthesis

There was no overall difference in the ratings provided by homosexual as compared to heterosexual listeners ($F(1, 176) = 0.06$, $p = 0.81$, E = 3.2, CI = [-0.6, 7.1]), nor in the ratings of male voices as compared to female voices ($F(1, 35.4) = 0.17$, $p = 0.69$, E = -4.3, CI = [-11.1, 2.5]). However, there was a large effect of voice modulation ($F(1, 17786) = 185.6$, $p < 0.001$) such that baseline voices were strongly preferred over any modulated voice (Masculinised: E = -6.9, CI = [-7.8, -6.0], Most masculinised: E = -21.4, CI = [-22.3, -20.5], Feminised: E = -13.4, CI = [-14.3, -12.5], Most feminised: E = -30.4, CI = [-31.1, -29.3]). See Figure 2a.

The attractiveness ratings given to voices of either sex were not different for heterosexual or homosexual listeners rating voices of their preferred sex ($F(1, 176) = 0.07$, $p = 0.79$, E = -0.5, CI = [-8.2, 7.2]). The overall preference for baseline over modulated voices was observed more strongly for homosexual than heterosexual listeners ($F(4, 17786) = 14.2$, $p < 0.001$; Masculinised: E = -0.9, CI = [-2.7, 0.9], Most Masculinesed: E = -3.3, CI = [-5.1, -1.5], Feminised: E = -3.6, CI = [-5.4, -1.8], Most feminised: E = -6.2, CI = [-8.0, -4.4]). Voice modulations affected ratings of male voices differently than female voices ($F(4, 17786) = 185.6$, $p < 0.001$), such that listeners were more tolerant of modulations that masculinised male voices relative to female voices, in line with gendered voice preferences (Masculinised: E = 10.5, CI = [8.7, 12.3], Most masculinised: E = 19.1, CI = [17.3, 20.9], Feminised: E = -1.9, CI = [-3.8, -0.1], Most feminised: E = 1.1, CI = [-0.7, 2.9]).

However, this preference for gender typical voice modulations (i.e., masculinised male voices and feminised female voices) was less apparent among homosexual listeners ($F(4, 17786) = 52.5$, $p < 0.001$) such that masculinised male voices were preferred less strongly and feminised male voices were disliked less strongly than expected from trends observed among heterosexual listeners (Masculinised: E = -8.1, CI = [-11.7, -4.5], Most masculinised: E = -11.8, CI = [-15.4, -8.2], Feminised: E = 8.9, CI = [5.3, 12.5], Most feminised: E = 8.5, CI = [4.9, 12.1]). These interactions demonstrate a pattern of voice preferences more complex than those captured by the preregistered hypotheses. Only heterosexual males on average preferred feminised voices, with all other listener groups preferring masculinised voices (see Figure 2b).

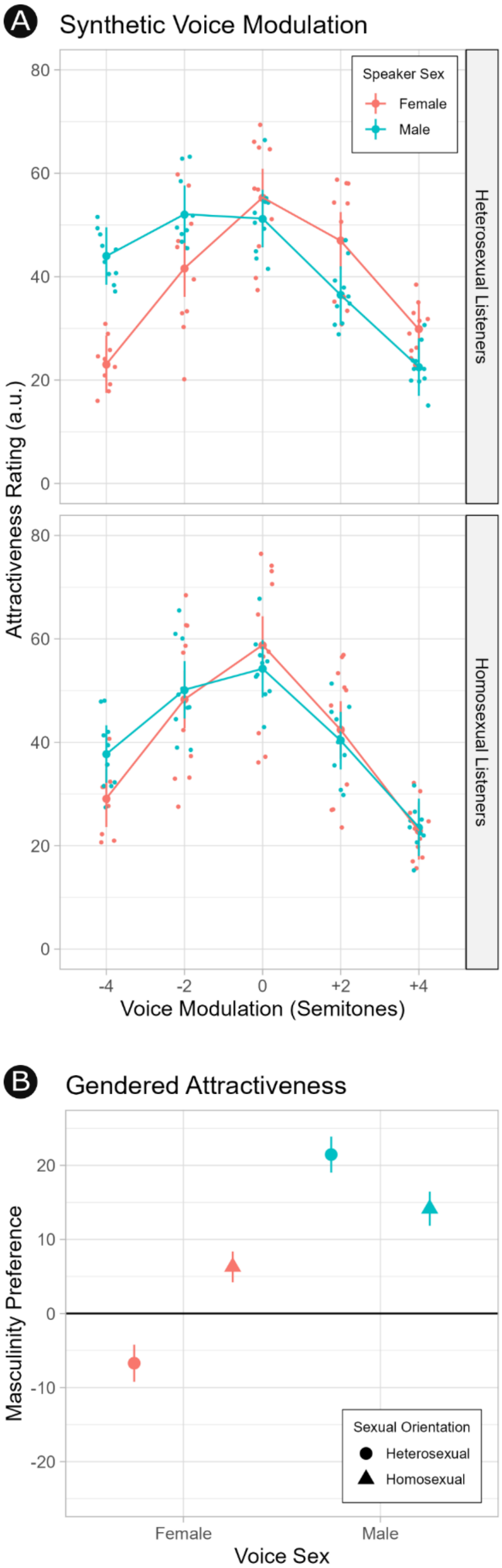

Figure 2: *Attractiveness ratings for synthetically modulated voices in Experiment 1. A) Connected points reflect estimates for each voice modulation condition, for female (salmon) or male (cyan) speakers as rated by heterosexual (upper panel) or homosexual (lower panel) listeners. Speakers strongly preferred unmodulated voices (0 Semitones). Heterosexual listeners were more forgiving of male voices that were masculinised by lowering f0 and lengthening VTL (-2, -4 Semitones) and female voices that were feminised by raising f0 and shortening VTL (+2, +4 semitones). Homosexual listeners (bottom) displayed a weaker preference for masculinised male voices and little evidence of a preference for feminised female voices. Small unconnected points are individual estimates for each speaker in the stimulus set to demonstrate inter-speaker variability. B) Differences in attractiveness ratings at the extremes of voice modulation calculated as the difference in attractiveness ratings given to stimuli modulated -4 semitones versus stimuli modulated +4 semitones. Values above 0 indicate a preference for masculinised voices.*

Experiment 2: Modulation by imitation

Results from Experiment 2 replicated some basic effects from Experiment 1 such as the preference for unmodulated voices and the greater sympathy for sex-typical voice modulation. However, effects were weaker and less consistent potentially due to greater variability in the performance of individual speakers at each vocal modulation step (See Figure 3).

There was no overall difference in the ratings provided by homosexual as compared to heterosexual listeners ($F(1, 176) = 0.06$, $p = 0.80$, E = -2.9, CI = [-6.1, 0.4]), nor in the ratings of male voices as compared to female voices ($F(1, 25.3) = 0.44$, $p = 0.51$, E = -0.03, CI = [-7.5, 7.5]). However, there was a large effect of voice modulation ($F(1, 17786) = 161.7$, $p < 0.001$) grossly consistent with the preference for baseline voices observed in Experiment 1 though less systematically (Masculinised: E = 1.1, CI = [0.2, 2.0], Most masculinised: E = -6.3, CI = [-7.3, -5.4], feminised: E = -7.8, CI = [-8.8, -6.9], Most feminised: E = -7.3, CI = [-8.3, -6.4]).

An interaction between voice sex and listener's sexual orientation indicated a tendency for heterosexual males to provide higher ratings than other groups ($F(1, 176) = 11.7$, $p < 0.001$, E = 11.9, CI = [5.4, 18.4]). In contrast to findings from Experiment 1, the overall preference for baseline over modulated voices was observed less strongly for homosexual than heterosexual listeners ($F(4, 17786) = 11.3$, $p < 0.001$; Masculinised: E = 3.0, CI = [1.1, 4.9], Most masculinised: E = 0.2, CI = [-1.6, 2.1], Feminised: E = 5.0, CI = [3.1, 6.8], Most feminised: E = 4.2, CI = [2.3, 6.0]). Voice modulations affected ratings of male voices differently than female voices ($F(4, 17786) = 67.2$, $p < 0.001$), such that extreme modulations of male voices were penalised more strongly (Masculinesed: E = -6.9, CI = [-8.8, -5.0], Most masculinised: E = 1.4, CI = [-0.5, 3.3], Feminised: E = 2.8, CI = [1.0, 4.7], Most feminised: E = -9.8, CI = [-11.6, -7.9]).

However, this preference for more moderate male voices was subject to idiosyncratic differences reflected by a complex interaction between voice sex, listener's sexual orientation, and voice modulation level ($F(4, 17786) = 15.6$, $p < 0.001$), for instance large masculinising modulations of male voices were rated highly by homosexual male listeners despite the broad tendency for aversion towards modulated voices (Masculinised: E = -1.8, CI = [-5.6, 1.9]; Most masculinised: E = 6.8, CI = [3.1, 10.5], Feminised: E = -6.9, CI = [-10.6, -3.1], Most feminised: E = -5.1, CI = [-8.8, -1.4]).

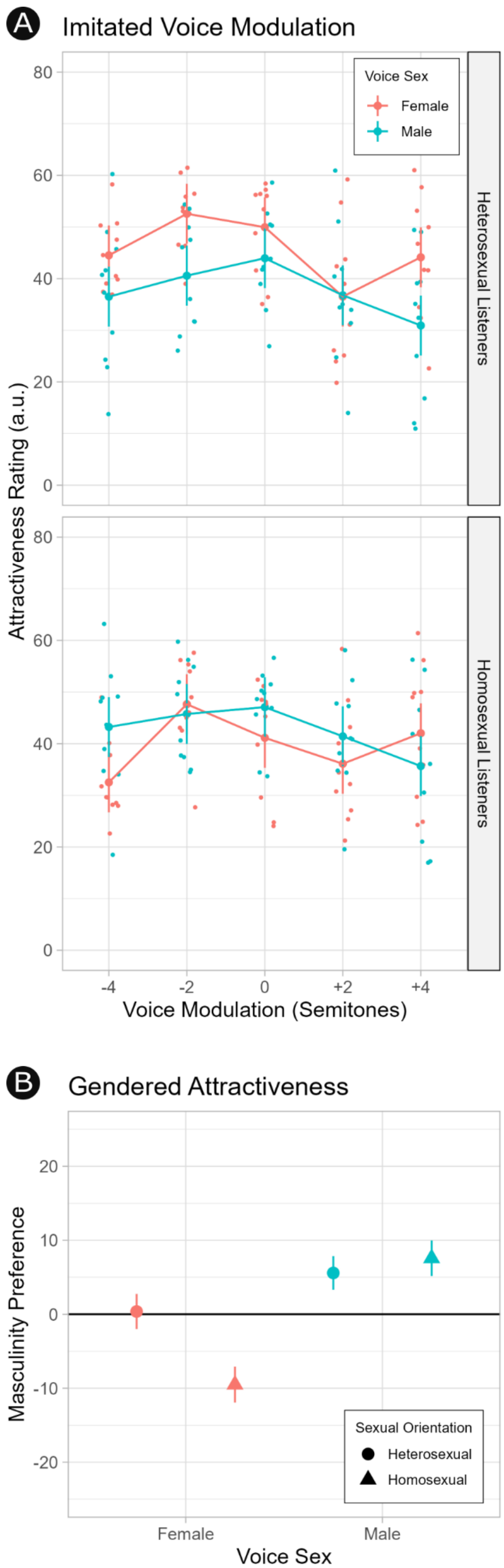

Figure 3: *Attractiveness ratings for voices modulated by imitation in Experiment 2. A) Connected points reflect estimates for each voice modulation condition, for female (salmon) or male (cyan) speakers as rated by heterosexual (top) or homosexual (bottom) listeners. Speakers strongly preferred unmodulated voices (0 Semitones). Listeners responded less consistently to voice modulations than demonstrated in experiment 1. Small unconnected points are individual estimates for each speaker in the stimulus set to demonstrate inter-speaker variability. Trends in the preferences of listeners may have been influenced by individual differences in the speakers' performances. B) Differences in attractiveness ratings at the extremes of voice modulation calculated as the difference in attractiveness ratings awarded to stimuli modulated -4 semitones versus stimuli modulated +4 semitones. Values above 0 indicate a preference for masculinised voices.*

Variability in homosexual voice preferences

A Levene's test for homogeneity of variances was conducted to test the hypothesis that homosexual listeners would be more varied in their voice attractiveness ratings than heterosexual listeners.

Levene's tests suggested inhomogeneity of variances for Experiment 1 ($F(19, 880)) = 3.4$, $p < 0.001$), but not Experiment 2 ($F(19, 880) = 0.7$, $p = 0.83$). However, an examination of the model residuals (see Figure 4) demonstrates no clear evidence that the attractiveness ratings of homosexual listeners were more varied than those of heterosexual raters. To the extent that LMMs are robust to violations of distributional assumptions (Schielzeth et al., 2020) it is not anticipated that the modest heteroscedasticity in Experiment 1 meaningfully impacted fixed effects estimates.

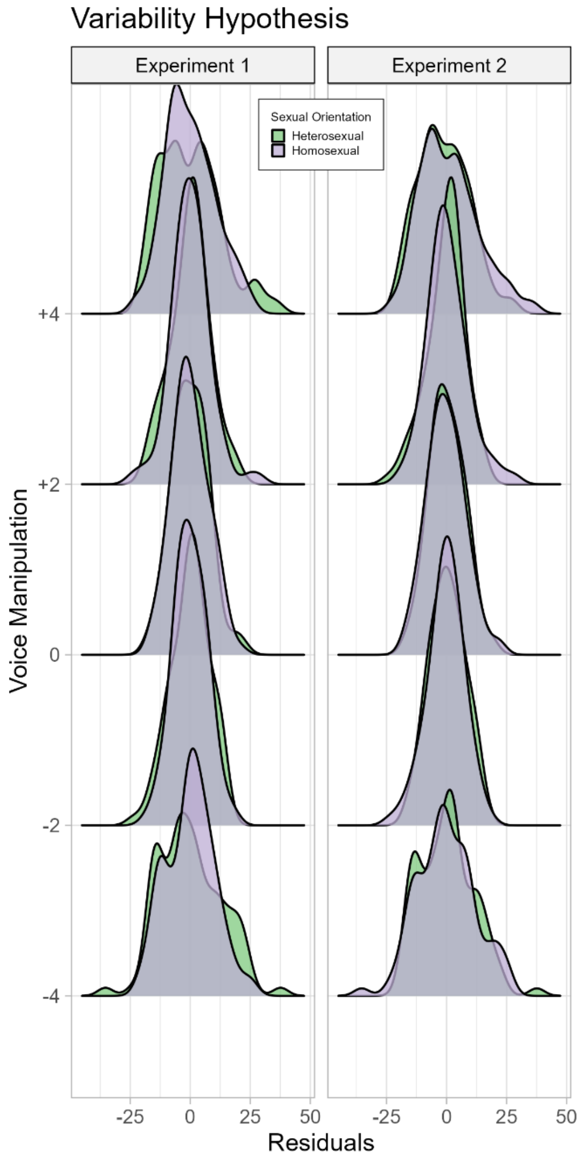

Figure 4: *Ridge plots showing distributions of model residuals. Mean attractiveness ratings were calcuated for each participant and voice modulation condition and refit by LMM. Residuals of this model reflect variability in attractiveness ratings. Wider distributions reflect greater individual differences. No evidence is observed that vocal attractiveness ratings by homosexual listeners were more variable than ratings by heterosexual listeners.*

Post-hoc comparison of synthetic versus imitated voice modulation

Mean attractiveness ratings were computed for each speaker at each voice modulation level in order to compare the attractiveness of voices that were modulated by digital synthesis versus volitional voice modulation (via imitation) (see Figure 5). At baseline, synthesised voices were perceived as more attractive. However with increasing degrees of modulation this preference shifted towards imitated voices ($F(4, 72) = 9.8$, $p < 0.001$; Masculinised: E = -8.0 , CI = [-15.5, -0.4], Most masculinised: E = -15.1, CI = [-22.6, -7.6], Feminised: E = -5.6, CI = [-13.1, 2.0], Most feminised: E = -22.9, CI = [-30.4, -15.4]). This effect has a similar magnitude in speakers selected for poor vocal modulation ability as speakers selected for good vocal modulation ability ($F(4, 72) = 0.6$, $p = 0.68$), Masculinised: E = -3.2, CI = [-18,2, 11.9], Most Masculinesed: E = -4.0, CI = [-19.1, 11.0], Feminised: E = -11.6, CI = [-26.7, 3.4], Most feminised: E = -4.9, CI = [-20.0, 10.1]), despite the larger changes to vocal acoustics that were produced by good vocal modulators.

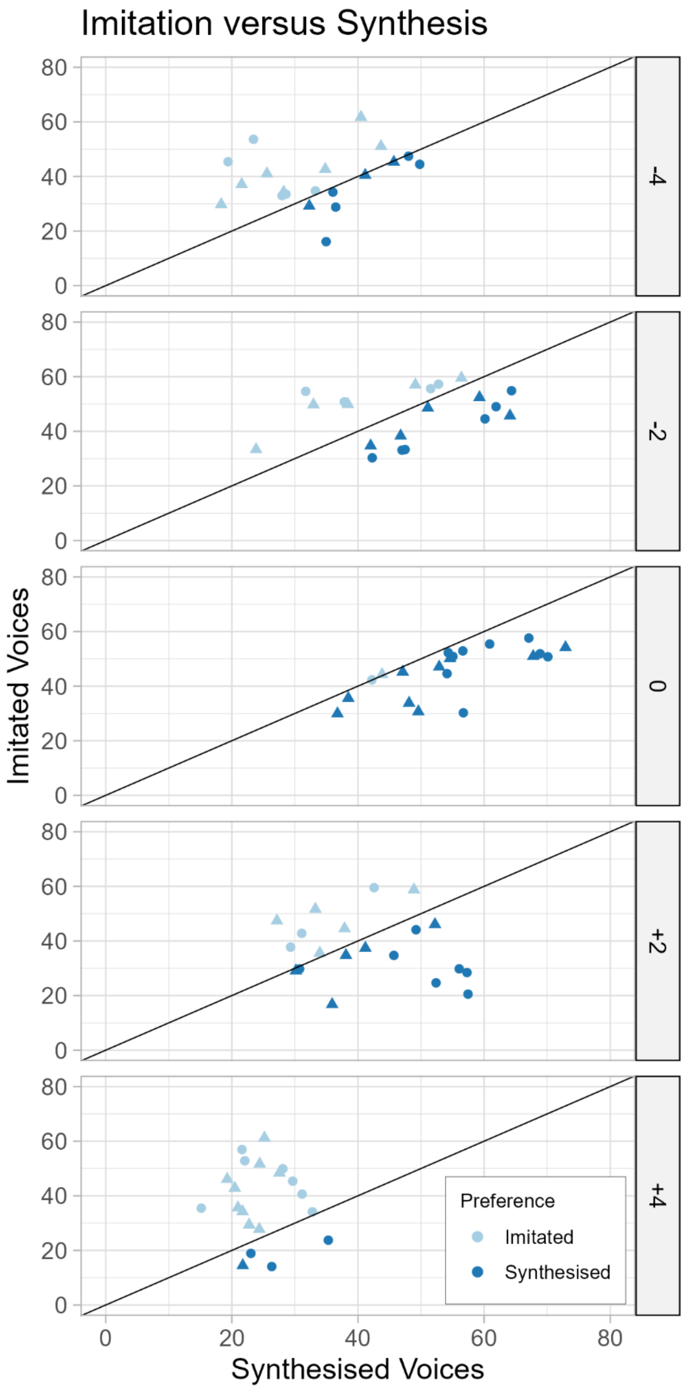

Figure 5: *Attractiveness ratings by individual speaker. Points indicate the mean attractiveness ratings for each speaker when their voice was modulated by synthesis (x-axis) or by imitation (y-axis), regardless of the gender or sexual orientation of the listener. Ratings are indicated separately for degrees of voice modulation (rows). The diagonal line indicates equal preference for synthesized (experiment 1) versus imitated (experiment 2) voices. Points below the diagonal (see row 0 Semitones) were preferred as synthesised voices. Points above the diagonal (see rows +/- 4 Semitones) were preferred as imitated voices. As the degree of voice modulation increases, preferences shift towards imitated voices. This pattern holds for both good (●) and for poor (▲) vocal modulators.*

Post-hoc comparison of speaker attractiveness at baseline

A post-hoc analysis of the mean attractiveness ratings for each speaker at baseline indicates that there were strong and stable differences between speakers even in the absence of modulation (see Figure 6). Linear regression confirmed that there was strong agreement between heterosexual and homosexual listeners. Voice modulation aside, there are large differences in the perceived attractiveness of different speakers and these individual differences are stable across sexual orientations.

In Experiment 1, the mean attractiveness of each speaker as rated by homosexual listeners was predicted by heterosexual ratings of the same stimulus for female speakers ($F(1, 16) = 257$, $p < 0.001$, $E = 1.35$, $CI = [1.16, 1.54]$), however a significant interaction indicated that the slope of this relation was shallower for male speakers ($F(1, 16) = 9.6$, $E = -0.52$, $CI = [-0.88, -0.17]$).

This pattern was replicated in the separate cohort of listeners recruited for Experiment 2. The mean attractiveness of each speaker to homosexual listeners was predicted by heterosexual ratings of the same stimulus for female speakers ($F(1, 16) = 156.3$, $p < 0.001$, $E = 1.2$, $CI = [1.0, 1.4]$), which again had a slightly shallower slope for male speakers ($F(1,16) = 12.9$, $p = 0.002$, $E = -0.46$, $CI = [-0.73, -0.19]$).

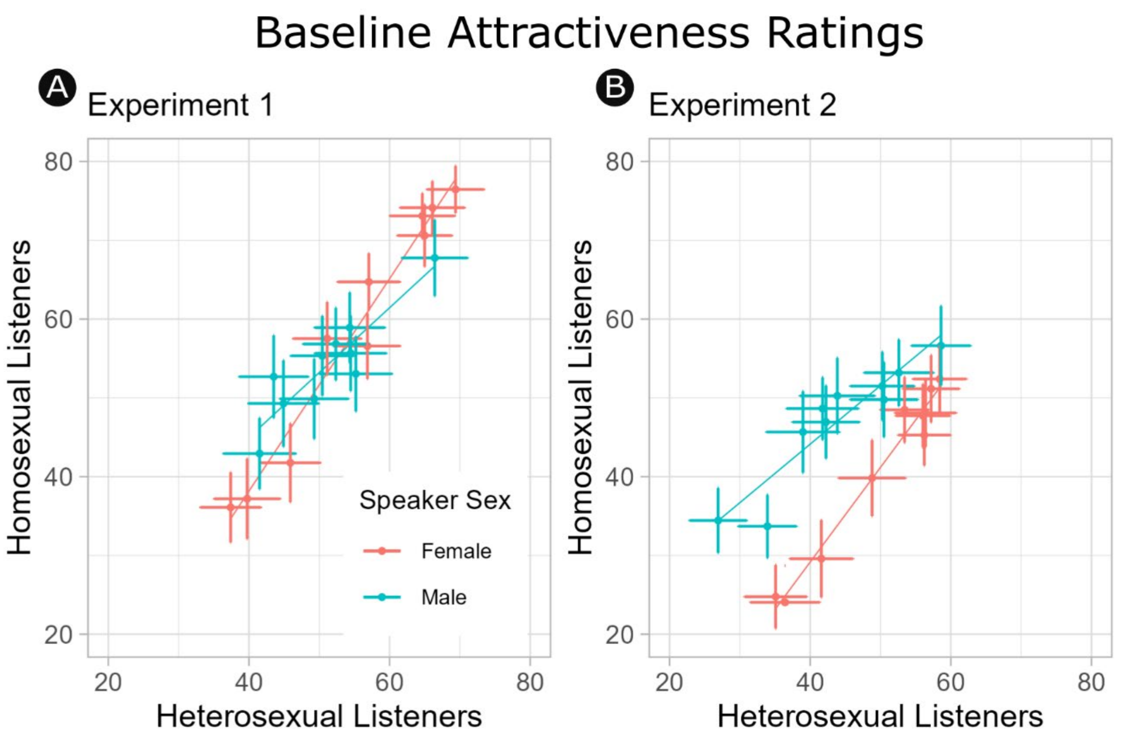

Figure 6: *Attractiveness ratings for each speaker at baseline in A) experiment 1 and B) experiment 2. Each point indicates the estimate for a given speaker pooling ratings across listeners. Scores on the x-axis indicate ratings of each speaker's baseline voice from heterosexual listeners while scores on the y-axis indicate ratings from homosexual listeners. Lines indicate 95% confidence intervals for either axis. The strong agreement across sexual orientations indicates robust differences between speakers independent from voice modulation.*

This shared pattern of preferences at baseline contradicts the observation that heterosexual males voice modulation preferences diverge from all other participant groups. A series of exploratory analyses from the low-level acoustic features of the stimuli including fundamental frequency, formant dispersion, harmonic-to-noise-ratio, and spectral slope did not adequately account for gross agreement between groups (see Supplementary Material 4). Further exploration is required to identify drivers of shared preferences for individual differences in speakers' voices that may exist alongside divergent preferences for voice modulations.

Discussion

Heterosexual listeners' attractiveness judgements replicated previous findings in that they preferred voice modulations that exaggerated sexually dimorphic vocal traits over voice modulations that diminished them. However, considering sexual orientation in our experimental design produced a more complex configuration that was not consistent with the listener hypothesis, the speaker hypothesis, or the variability hypothesis specified at study preregistration. Instead, we observed a general preference for masculinising voice modulations, to which heterosexual male listeners the only consistent exception. This effect was particularly evident in Experiment 1, in which stimuli were more controlled, with moderate (+/- 2 semitone) voice modulation yielding weaker but qualitatively similar effects to stronger voice modulation (+/- 4 semitones). In Experiment 2, where stimuli were more naturalistic but less controlled, there was less agreement between moderate and strong voice modulations which may reflect idiosyncrasies in the performances of individual speakers.

Masculinity bias

These findings reflect a tension between the use of the voice for biologically and socially oriented signalling. The voice serves as a fitness signal that listeners use to assess, and speakers use to exaggerate, secondary sexual characteristics. Simultaneously, it also serves as a social signal that is perceived as a signal of confidence, competence, and relative position within social hierarchies.

Feinberg (2008) argued that both faces and voices can act as a cue to mate quality. Indeed there is a considerable body of evidence that facial and vocal attractiveness are correlated in both females (Feinberg et al., 2005; Fraccaro et al., 2010; Lander, 2008; Valentova et al., 2017; Wells et al., 2013) and males (Feinberg et al., 2008; Rezlescu et al., 2015; Saxton et al., 2006). The most likely mechanism linking these signals are sex hormones such as testosterone that have a broad influence over physical development, including the morphology of the vocal tract (Markova et al., 2016). However, the voice is a highly dynamic and multipurposed communicative channel. Cues to attractiveness compete with other contextually important signals (e.g., signalling sexual interest) which may pull voice modulation in competing directions (Hughes et al., 2010, 2014; Pisanski et al., 2018). Listeners also readily make a range of social trait judgements from even very brief exposure to voices (Guldner et al., 2024; Lavan, 2023; Mahrholz et al., 2018; McAleer et al., 2014; Mileva & Lavan, 2023; Sorokowski et al., 2023). These judgements can be adequately summarised by components of valence (correlated to trustworthiness, warmth and likability) and social dominance (correlated to confidence and aggressiveness). These trait judgements are correlated with f0 and formant dispersion among other acoustical cues, but in different combinations for female and male voices. As a result, perceived social traits complicate a strictly biological interpretation of vocal attractiveness.

The masculine voice bias observed by the present study is indicative of the competing demands on the female voice in enhancing status as a secondary sexual characteristic versus as a cue to the speaker's position within social hierarchies. Whereas male speakers may be able to adopt a consistent vocal strategy for inducing favourable judgements from listeners, (heterosexual) female speakers may have to select among competing priorities.

The naturalism preference

Listeners disliked modulated voices regardless of sex or sexual orientation. Indeed, even the additive effect of sexually dimorphic voice preferences manifested as increased tolerance for modulation alongside a stronger preference for unmodulated baseline voices.

This antipathy towards modulated voices may reflect the conflicting interests of speakers and listeners. Unlike the contents of spoken language in which speakers exchange information through propositional statements, most systems of communication that have been observed in nature are competitive – the interests of the signaller are not necessarily aligned with the interests of the receiver (Mokkonen & Lindstedt, 2016; Scott-Phillips, 2008). In the specific case of signals to attractiveness, the signaller has a vested interest in over-representing their potential value as a mate, or as a partner, whereas the receiver has a vested interest in making an accurate valuation. Moreover, how much weight to place on any given signal, if any at all, is a decision that rests in the ear of the beholder. It is a principle of animal communication that signalling and cueing systems are only retained by natural selection when they are on average honest, as otherwise the intended receiver of the signal may not be inclined to listen (Maynard Smith & Harper, 2003; Searcy & Nowicki, 2005).

Vocal size exaggeration is evident across a broad range of terrestrial mammals (Charlton & Reby, 2016). A classic example is found in the roars of Red Deer (*Cervus elaphus*) which are produced by males during the mating season. Male Red Deer roar while lowering their larynx, effectively exaggerating their apparent vocal size. Males use these roars as part of intra-sexual agonistic contests (Reby et al., 2005), while females use them to inform mate choice decisions (Charlton et al., 2007). Though the presence of vocal size exaggeration makes these signals potentially dishonest, the degree of exaggeration is constrained by the signallers' vocal anatomy; bigger animals can exaggerate more (Reby & McComb, 2003). Hence, the roar of the red deer is an honest indicator of relative, if not absolute, body size. Similar principles constrain human vocalisations, keeping them relatively honest (Aung et al., 2021; Kleisner et al., 2021; Raine et al., 2018, 2019).

We previously observed that male speakers were constrained by the range of motion available for exaggerating vocal size while female speakers were constrained by the range of motion available for attenuating vocal size (Belyk et al., 2022), demonstrating that anatomical constraints may work to keep vocal size cues at least partially honest in humans. Nonetheless, the present findings provide further evidence that listeners may attempt to identify instances of vocal modulation as a defence against manipulative signalling (Pisanski & Reby, 2021).

Voice versus voice modulation

Despite the diverging voice modulation preferences of heterosexual males from the other participant groups, there was strong consensus on which speakers sounded most attractive at baseline. This pattern may be indicative of global voice preferences separate from those manipulated by vocal modulation. For instance, analysis of the acoustical features of these stimuli did not reveal a clear relationship with either the fundamental frequency or apparent vocal size cues that were the basis of vocal modulation in this study.

We explored apparent vocal health as an alternative mechanism that could drive such global agreement. As with other trait judgements, listeners will readily express an opinion about the health of a speaker with high inter-rater agreement emerging even with relatively brief exposures (Lavan, 2023; Sorokowski et al., 2024). Some previous evidence suggests that listeners may conflate health with attractiveness (Albert et al., 2021). However, an exploratory analysis of two acoustical cues with plausible relationships to vocal health yielded only weak effects that do not adequately account for the strong consensus observed across speakers.

Regardless, the data do clearly demonstrate that preferences for voices are not necessarily the same as preferences for voice modulations. These two related but distinct channels of information may reflect separate drivers of behaviour based on biological signals carrying attributes of the body to which the voice belongs and social signals communicating the attitudes and intentions of the speaker.

Inclusive voice science

Previous research in this area has focused disproportionately on the perceptions of heterosexual listeners (see notable exception in Valentová et al., 2013). While some previous research suggests that vocal characteristics may differ as a function of sexual orientation (Baeck et al., 2011), these are more likely to be flexible or cultural signals rather than bioacoustic markers of sexual orientation (Daniele et al., 2020) and pervasive beliefs to the contrary may contribute to stigmatisation (Fasoli et al., 2021). While sexual orientation is not listed among the Western, Educated, Industrialised, Rich, and Democratic (WEIRD) populations that are over-sampled by psychological science (Henrich et al., 2010), the risks of overgeneralising from demographically narrow samples is similar. Far from merely being in agreement, or not, with patterns observed in heterosexual listeners, sampling a broader cross section of the population revealed a richer context in which to understand human vocal communication. Further research should extend inclusion to a more diverse range of sexual orientations.

Conclusions

We replicated previous findings that heterosexual listeners find voice modulations that exaggerate secondary sexual characteristics more attractive than voice modulations that diminish them. However, inclusion of homosexual listeners revealed that the heterosexual male preference for feminised voices may be an exception to a generalised bias towards masculinised voices. In addition to the influence of voice modulation, there were global preferences for unmodulated voices over modulated voices, and for naturalistic voice modulations over synthetic ones that are consistent with listeners attempts to defend against potentially manipulative vocal signals. Finally, among unmodulated voices there was strong agreement within and between groups about which speakers' voices were most attractive consistent with coexisting preferences for voices alongside voice modulations as cues to the biology of the individual and their social intentions, respectively. Widening the inclusivity of research provided a richer context in which to study the interacting influences of sex, sexual orientation, and voice modulation on listener's perceptions of vocal attractiveness.

Open Practices

Processed data, a data dictionary, and analysis code are archived at https://osf.io/ywpqc/?view_only=7721fad22c0b42e4852bcdcf8e69e4ab. Study preregistrations for experiment 1 and experiment 2 are available at https://osf.io/w63au/?view_only=70d036cb7441447bb2bfe1f897c60e30 and https://osf.io/h8cg7/?view_only=8551a45b9e2f4e3eabfceea8c4b19c3d, respectively.

Funding

Carolyn McGettigan is funded by a Research Leadership Award (RL-2016-013) from The Leverhulme Trust.

Acknowledgements

We thank Camilla Beone for her contributions to data collection.

References

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, *13*(3), 424–446. https://doi.org/10.1016/S0892-1997(99)80048-4

Albert, G., Arnocky, S., Puts, D. A., & Hodges-Simeon, C. R. (2021). Can listeners assess men's self-reported health from their voice? *Evolution and Human Behavior*, *42*(2), 91–103. https://doi.org/10.1016/j.evolhumbehav.2020.08.001

Anikin, A. (2019). Soundgen: An open-source tool for synthesizing nonverbal vocalizations. *Behavior Research Methods*, *51*(2), 778–792. https://doi.org/10.3758/s13428-018-1095-7

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*(1659), 1077–1082. https://doi.org/10.1098/rspb.2008.1542

Aung, T., Goetz, S., Adams, J., McKenna, C., Hess, C., Roytman, S., Cheng, J. T., Zilioli, S., & Puts, D. (2021). Low fundamental and formant frequencies predict fighting ability among male mixed martial arts fighters. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-020-79408-6

Babel, M., McGuire, G., & King, J. (2014). Towards a more nuanced view of vocal attractiveness. *PLoS ONE*, *9*(2), 1–10. https://doi.org/10.1371/journal.pone.0088616

Baeck, H., Corthals, P., & Borsel, J. V. (2011). Pitch characteristics of homosexual males. *Journal of Voice*, *25*(5), e211–e214. https://doi.org/10.1016/j.jvoice.2010.10.019

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belyk, M., & Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience & Biobehavioral Reviews*, *77*, 177–193. https://doi.org/10.1016/j.neubiorev.2017.03.014

Belyk, M., Waters, S., Kanber, E., Miquel, M. E., & McGettigan, C. (2022). Individual differences in vocal size exaggeration. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-05170-6

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* (6.3.17) [Computer software]. http://www.praat.org/

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Brunings, J. W., Schepens, J. J. B. F. G., Peutz-Kootstra, C. J., & Kross, K. W. (2013). The expression of estrogen and progesterone receptors in the human larynx. *Journal of Voice: Official Journal of the Voice Foundation*, *27*(3), 376–380. https://doi.org/10.1016/j.jvoice.2013.01.011

Charlton, B. D., & Reby, D. (2016). The evolution of acoustic size exaggeration in terrestrial mammals. *Nature Communications*, *7*, 1–8. https://doi.org/10.1038/ncomms12739

Charlton, B. D., Reby, D., & McComb, K. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*(4), 382–385. https://doi.org/10.1098/rsbl.2007.0244

Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*, 790–793.

Daniele, M., Fasoli, F., Antonio, R., Sulpizio, S., & Maass, A. (2020). Gay voice: Stable marker of sexual orientation or flexible communication device? *Archives of Sexual Behavior*, *49*(7), 2585–2600. https://doi.org/10.1007/s10508-020-01771-2

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*(4), 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033

Fasoli, F., Hegarty, P., & Frost, D. M. (2021). Stigmatization of ‘gay-sounding’ voices: The role of heterosexual, lesbian, and gay individuals’ essentialist beliefs. *British Journal of Social Psychology*, *60*(3), 826–850. https://doi.org/10.1111/bjso.12442

Feinberg, D., Jones, B., Little, A., Burt, D., & Perrett, D. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

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. https://doi.org/10.1002/evan.20166

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Little, A. C. (2008). Correlated preferences for men’s facial and vocal masculinity. *Evolution and Human Behavior*, *29*(4), 233–241. https://doi.org/10.1016/j.evolhumbehav.2007.12.008

Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Law Smith, M. J., Cornwell, R. E., Tiddeman, B. P., Boothroyd, L. G., & Perrett, D. I. (2005). The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, *26*(5), 398–408. https://doi.org/10.1016/j.evolhumbehav.2005.04.001

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 1511–1522. https://doi.org/10.1121/1.427148

Fraccaro, P. J., Feinberg, D. R., DeBruine, L. M., Little, A. C., Watkins, C. D., & Jones, B. C. (2010). Correlated male preferences for femininity in female faces and voices. *Evolutionary Psychology*, *8*(3), 147470491000800311. https://doi.org/10.1177/147470491000800311

Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., Little, A. C., & Debruine, L. M. (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

Fraccaro, P. J., O'Connor, J. J. M., Re, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, *85*(1), 127–136. https://doi.org/10.1016/j.anbehav.2012.10.016

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, *27*(15), 2865–2873. https://doi.org/10.1002/sim.3107

Gelman, A., & Yo, Y.-S. (2015). *arm: Data analysis using regression and multilevel/hierarchical models* [Computer software].

Glassenberg, A. N., Feinberg, D. R., Jones, B. C., Little, A. C., & DeBruine, L. M. (2010). Sex-dimorphic face shape preference in heterosexual and homosexual men and women. *Archives of Sexual Behavior*, *39*(6), 1289–1296. https://doi.org/10.1007/s10508-009-9559-6

Groyecka, A., Pisanski, K., Sorokowska, A., Havlíček, J., Karwowski, M., Puts, D., Roberts, S. C., & Sorokowski, P. (2017). Attractiveness is multimodal: Beauty is also in the nose and ear of the beholder. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00778

Guldner, S., Lavan, N., Lally, C., Wittmann, L., Nees, F., Flor, H., & McGettigan, C. (2024). Human talkers change their voices to elicit specific trait percepts. *Psychonomic Bulletin & Review*, *31*(1), 209–222. https://doi.org/10.3758/s13423-023-02333-y

Harries, M. L. L., Walker, J. M., Williams, D. M., Hawkins, S., & Hughes, I. A. (1997). Changes in the male voice at puberty. *Archives of Disease in Childhood*, *77*(5), 445–447. https://doi.org/10.1136/adc.77.5.445

Henrich, J., Heine, J. H., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*, 29.

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. https://doi.org/10.1007/s10919-010-0087-9

Hughes, S. M., Harrison, M. A., & Gallup, G. G. (2002). The sound of symmetry: Voice as a marker of developmental instability. *Evolution and Human Behavior*, *23*(3), 173–180. https://doi.org/10.1016/S1090-5138(01)00099-X

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

Kahane, J. C. (1978). A morphological study of the human prepubertal and pubertal larynx. *American Journal of Anatomy*, *151*(1), 11–19. https://doi.org/10.1002/aja.1001510103

Kleisner, K., Leongómez, J. D., Pisanski, K. K., Fiala, V., Cornec, C., Groyecka-Bernard, A., Butovskaya, M., Reby, D., Sorokowski, P., & Akoko, R. M. (2021). Predicting strength from aggressive vocalizations versus speech in African bushland and urban communities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1840). https://doi.org/10.1098/rstb.2020.0403

Lander, K. (2008). Relating visual and vocal attractiveness for moving and static faces. *Animal Behaviour*, *75*(3), 817–822. https://doi.org/10.1016/j.anbehav.2007.07.001

Lavan, N. (2023). The time course of person perception from voices: A behavioral study. *Psychological Science*, *34*(7), 771–783. https://doi.org/10.1177/09567976231161565

Mahrholz, G., Belin, P., & McAleer, P. (2018). Judgements of a speaker's personality are correlated across differing content and stimulus type. *PLoS ONE*, *13*(10), 1–22. https://doi.org/10.1371/journal.pone.0204991

Markova, D., Richer, L., Pangelinan, M., Schwartz, D. H., Leonard, G., Perron, M., Pike, G. B., Veillette, S., Chakravarty, M. M., Pausova, Z., & Paus, T. (2016). Age-and sex-related variations in vocal-tract morphology and voice acoustics during adolescence. *Hormones and Behavior*, *81*, 84–96. https://doi.org/10.1016/j.yhbeh.2016.03.001

Maynard Smith, J., & Harper, D. (2003). *Animal Signals*. Oxford University Press.

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'hello'? Personality impressions from brief novel voices. *PLoS ONE*, *9*(3), 1–9. https://doi.org/10.1371/journal.pone.0090779

Mileva, M., & Lavan, N. (2023). Trait impressions from voices are formed rapidly within 400 ms of exposure. *Journal of Experimental Psychology: General*, *152*(6), 1539–1550. https://doi.org/10.1037/xge0001325

Milne, A. E., Bianco, R., Poole, K. C., Zhao, S., Oxenham, A. J., Billig, A. J., & Chait, M. (2020). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *2017*. https://doi.org/10.3758/s13428-020-01514-0

Mokkonen, M., & Lindstedt, C. (2016). The evolutionary ecology of deception. *Biological Reviews*, *91*(4), 1020–1035. https://doi.org/10.1111/brv.12208

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*(1), 72–81. https://doi.org/10.1016/S0892-1997(00)80096-X

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, *20*, 304–318. https://doi.org/10.1016/j.tics.2016.01.002

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M. M., Röder, S., Andrews, P. W., Fink, B., DeBruine, L. M., Jones, B. C., & Feinberg, D. R. (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., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., & Reby, D. (2018). Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1893). https://doi.org/10.1098/rspb.2018.1634

Pisanski, K., & Reby, D. (2021). Efficacy in deceptive vocal exaggeration of human body size. *Nature Communications*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-21008-7

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, *129*(4), 2201–2212. https://doi.org/10.1121/1.3552866

R Core Team. (2019). *R: A language and environment for statistical computing.* [Computer software]. R Foundation for Statistical Computing. http://cran.r-project.org/

Raine, J., Pisanski, K., Bond, R., Simner, J., & Reby, D. (2019). Human roars communicate upper-body strength more effectively than do screams or aggressive and distressed speech. *PLOS ONE*, *14*(3), e0213034. https://doi.org/10.1371/journal.pone.0213034

Raine, J., Pisanski, K., Oleszkiewicz, A., Simner, J., & Reby, D. (2018). Human listeners can accurately judge strength and height relative to self from aggressive roars and speech. *iScience*, *4*, 273–280. https://doi.org/10.1016/j.isci.2018.05.002

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

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal: Biological Sciences*, *272*(1566), 941–947.

Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant voices and attractive faces: The contribution of visual and auditory information to integrated person impressions. *Journal of Nonverbal Behavior*, *39*(4), 355–370. https://doi.org/10.1007/s10919-015-0214-8

Saxton, T. K., Caryl, P. G., & Craig Roberts, S. (2006). Vocal and facial attractiveness judgments of children, adolescents and adults: The ontogeny of mate choice. *Ethology*, *112*(12), 1179–1185. https://doi.org/10.1111/j.1439-0310.2006.01278.x

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, *11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434

Scott-Phillips, T. C. (2008). Defining Biological Communication. *Journal of Evolutionary Biology*, *21*(2), 387–395.

Searcy, W. A., & Nowicki, S. (2005). *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems*. Princeton University Press.

Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., & Havlíček, J. (2017). Voices of Africa: Acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, *127*, 205–211. https://doi.org/10.1016/j.anbehav.2017.03.014

Shiramizu, V., Docherty, C., DeBruine, L. M., & Jones, B. C. (2020). Sexual orientation predicts men's preferences for sexually dimorphic face-shape characteristics: A replication study. *PLoS ONE*, *15*(11 November), 6–11. https://doi.org/10.1371/journal.pone.0242262

Sorokowski, P., Groyecka-Bernard, A., Frackowiak, T., Kobylarek, A., Kupczyk, P., Sorokowska, A., Misiak, M., Oleszkiewicz, A., Bugaj, K., Włodarczyk, M., & Pisanski, K. (2023). Comparing accuracy in voice-based assessments of biological speaker traits across speech types. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-49596-y

Sorokowski, P., Pisanski, K., Frąckowiak, T., Kobylarek, A., & Groyecka-Bernard, A. (2024). Voice-based judgments of sex, height, weight, attractiveness, health, and psychological traits based on free speech versus scripted speech. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02445-5

Story, B. H., & Titze, I. R. (1995). Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America*, *97*(2), 1249–1260. https://doi.org/10.1121/1.412234

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America*, *123*(5), 2733. https://doi.org/10.1121/1.2832337

Valentová, J., Roberts, S. C., & Havlíček, J. (2013). Preferences for facial and vocal masculinity in homosexual men: The role of relationship status, sexual restrictiveness, and self-perceived masculinity. *Perception*, *42*(2), 187–197. https://doi.org/10.1068/p6909

Valentova, J. V., Varella, M. A. C., Havlíček, J., & Kleisner, K. (2017). Positive association between vocal and facial attractiveness in women but not in men: A cross-cultural study. *Behavioural Processes*, *135*, 95–100. https://doi.org/10.1016/j.beproc.2016.12.005

Voelter, Ch., Kleinsasser, N., Joa, P., Nowack, I., Martínez, R., Hagen, R., & Voelker, H. U. (2008). Detection of hormone receptors in the human vocal fold. *European Archives of Oto-Rhino-Laryngology*, *265*(10), 1239–1244. https://doi.org/10.1007/s00405-008-0632-x

Waters, S., Kanber, E., Lavan, N., Belyk, M., Carey, D., Cartei, V., Lally, C., Miquel, M., & Mcgettigan, C. (2021). Singers show enhanced performance and neural representation of vocal imitation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, 20200399. https://doi.org/10.1098/rstb.2020.0399

Wells, T., Baguley, T., Sergeant, M., & Dunn, A. (2013). Perceptions of human attractiveness comprising face and voice cues. *Archives of Sexual Behavior*, *42*(5), 805–811. https://doi.org/10.1007/s10508-012-0054-0

Xu, Y., Lee, A., Wu, W. L., Liu, X., & Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0062397

Zamponi, V., Mazzilli, R., Mazzilli, F., & Fantini, M. (2021). Effect of sex hormones on human voice physiology: From childhood to senescence. *Hormones*, *20*(4), 691–696. https://doi.org/10.1007/s42000-021-00298-y