Facial Visualizations of Women’s Voices Suggest a Cross-Modality Preference for Femininity

Susanne Röder, Department of Sociobiology/Anthropology, University of Göttingen, Göttingen, Germany.

Bernhard Fink, Department of Sociobiology/Anthropology, University of Göttingen, Göttingen, Germany. Email: bernhard.fink@ieee.org (Corresponding author).

David R. Feinberg, Department of Psychology, Neuroscience and Behavior, McMaster University, Hamilton, Ontario, Canada.

Nick Neave, Department of Psychology, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK.

Abstract: Women with higher-pitched voices and more feminine facial features are commonly judged as being more attractive than are women with lower-pitched voices and less feminine faces, possibly because both features are affected by (age-related) variations in endocrine status. These results are primarily derived from investigations of perceptions of variations in single-modality stimuli (i.e., faces or voices) in samples of young adult women. In the present study we sought to test whether male and female perceptions of women’s voices affect visual representations of facial femininity. Eighty men and women judged voice recordings of 10 young girls (11-15 years), 10 adult women (19-28 years) and 10 peri-/post-menopausal women (50-64 years) on age, attractiveness, and femininity. Another 80 men and women were asked to indicate the face they think each voice corresponded to using a video that gradually changed from a masculine looking male face into a feminine looking female face. Both male and female participants perceived voices of young girls and adult women to be significantly younger, more attractive and feminine than those of peri-/post-menopausal women. Hearing young girls’ and adult women’s voices resulted in both men and women selecting faces that differed markedly in apparent femininity from those associated with peri-/post-menopausal women’s voices. Voices of young girls had the strongest effect on visualizations of facial femininity. Our results suggest a cross-modal preference for women’s vocal and facial femininity, which depends on female age and is independent of the perceiver’s sex.

Keywords: age, attractiveness, face, femininity, perception, voice, women
Facial visualizations of women’s voices

Introduction

Previous research has shown that men are remarkably sensitive to variations in women’s facial morphology and voice quality, both of which are linked to female age and fertility (see for review Feinberg, 2008; Puts, Jones, and DeBruine, 2012). Evolutionary psychologists, therefore, propose that preferences for certain female facial and vocal characteristics have evolved through sexual selection pressures operating on male perceptions in order to identify high quality mates. According to this logic, these preferences are expressed in the form of positive attractiveness perceptions, which is thought to reflect men’s assessments of key features that play a role in mate selection (for review see Fink and Penton-Voak, 2002; Gangestad and Scheyd, 2005).

Studies on men’s perceptions of women’s facial and vocal characteristics (among others) have repeatedly demonstrated that attractiveness judgments are strongly correlated with age (e.g., Buss, 1989; Rhodes, 2006) and femininity assessments (e.g., Jones et al., 1995; Perrett, May, and Yoshikawa, 1994;), possibly because both features are associated with estrogen levels (Abitbol, Abitbol, and Abitbol, 1999; Feinberg, DeBruine, Jones, and Perrett, 2008; Law-Smith et al., 2006), and thus reflect reproductive potential. Facial features that correspond to high estrogen levels include a small chin, high cheekbones, full lips, and large round eyes (Johnston and Franklin, 1993; Jones et al., 1995; Schaefer et al., 2006). Vocal fundamental frequency, which is perceived as voice pitch, is a sexually dimorphic feature with the average woman’s voice pitch being around twice as high as that of the average man (Titze, 1994). This sex difference seems to be strongly associated with the differential production of the sex steroids testosterone and estrogen during the pubertal years (Abitbol et al., 1999); the larynx being particularly sensitive to the variation in these hormones (Caruso et al., 2000). Voices of older people differ from those of younger people (Orlikoff, 1990; Ryan, 1972; Sweeting and Baken, 1982; Zenker, 1964), and voice pitch tends to decrease with increasing maturity, being positively related to changes in estrogen levels in women (Abitbol et al., 1999; Feinberg et al., 2006). Indeed, women’s vocal attractiveness is negatively associated with perceived age (Collins and Missing, 2003; Feinberg et al., 2008).

Thus, several studies have demonstrated positive associations between voice pitch and men’s attractiveness ratings of natural women’s voices (Collins and Missing, 2003; Feinberg et al., 2008; Fraccaro et al., 2011; Puts, Barndt, Welling, Dawood, and Burriss, 2011). Furthermore, other studies have systematically manipulated recordings of women’s voices by altering voice pitch, thus creating feminized and masculinized versions of them. Collectively, these studies have found in Canada, the US, UK, and among the Hadza (a hunter-gatherer group in Tanzania), that men judged the feminized versions of voices as more attractive than masculinized versions (Apicella and Feinberg, 2009; Feinberg et al., 2008; Jones, Feinberg, DeBruine, Little, and Vukovic, 2008, 2010; Klofstad et al., 2012; O’Connor et al., 2011; Re, O’Connor, Bennett, and Feinberg, 2012).

In women, facial and body morphology and voice quality seem to be correlated across modalities (Feinberg et al., 2005; Grammer, Fink, Juette, Ronzal, and Thornhill, 2001). For example, women whose faces were rated as attractive also received higher attractiveness judgments of their bodies (Thornhill and Grammer, 1999). Similarly, Collins
and Missing (2003) reported that women with attractive faces tend to also have attractive voices. These authors found that digital facial photographs of women with high-pitched voices were perceived to be more feminine and younger than photos presented with low-pitched voices. Therefore, male preferences for multiple traits like visual and vocal femininity are thought to be adaptive, since both indicate health and youth (Bryant and Haselton, 2009; Vukovic, Feinberg, DeBruine, Smith, and Jones, 2010). Researchers propose that multiple ornaments may have evolved, in part, to reduce the risk of being deceived when relying on a single cue or signal (Møller and Pomiankowski, 1993). The assumption here is that facial and vocal characteristics reflect the same underlying trait, i.e., sex-hormone levels (Feinberg, 2008), and the consistency in preferences for these traits may reduce the effects of noise and error perceptions of the qualities (see also Grammer et al., 2001).

Although many studies have investigated the relationship between vocal and facial attractiveness (e.g., Collins and Missing 2003; Feinberg et al., 2008), most research has mainly focused on attractiveness perception of faces or voices presented in isolation, and thus considered only single modality stimuli. Only one study to date has presented voices and faces manipulated in masculinity simultaneously (O’Connor et al., 2011) and found no interactions among vocal and facial attractiveness ratings. With regard to the wide range of variations in females’ voice pitch, affected by age-related hormonal changes during a woman’s life span (for review see Amir and Biron-Shental, 2004), we hypothesized that listeners isolate femininity and age-related variation out of unfamiliar female voices and translate this information into a facial visualization. Related to cross-modality preferences for female faces and voices, we expected (i) age-related judgments of voice recordings, with those of younger women being judged as younger, more attractive and feminine, and (ii) facial visualizations of the sound of younger women to appear more feminine.

Materials and Methods

Participants

Our sample was 30 female participants of three age groups, i.e., 10 young girls (11-15 years, \(M = 14.30; SD = 1.25\)), 10 adult women (19-28 years, \(M = 23.40; SD = 2.59\)) and 10 peri-/post-menopausal women (50-64 years, \(M = 55.80; SD = 4.85\)), who were recruited from a larger scale project on female sexual signals. All claimed to be German native speakers, non-smokers, and not having a cold or facial injury (e.g., a broken nose) that could have affected their individual voice quality. None reported to be taking any kind of hormonal medication (e.g., hormonal contraceptives or hormonal supplements).

Voice recordings

Voices were recorded with a unidirectional microphone (Rode SM6, with phantom power and pop filter), positioned approximately 10 centimeters in front of the mouth, using computer software (Apple Logic Studio®), in mono and at a sampling rate of 44.1 kHz with 16-bit amplitude quantization. We used a digital interface (M-AUDIO 8x8 Audio/MIDI-Interface Fast Track®) to encode the sounds. Female participants were requested to speak the five vowels (A [a], E [ɛ], I [i], O [ɔ], U [ʊ]) repeatedly and for 1 minute. To
Facial visualizations of women’s voices

ensure a constant speech rate, the vowels were presented visually on a computer screen in front of them (via a video clip), one after the other, and in 2-second intervals. Participants were asked to speak each vowel at the time it was presented on the screen. The serial order of the visual presentation of vowels changed six times within 1 minute of presentation in order to avoid habituation. For the subsequent rating study, a sequence of some 9 seconds of the vowels I [i], E [ɛ], O [ɔ], U [ʊ], A [a] (in that order), was digitally cropped out of the entire stream and saved in MP3 audio format (Constant Bit Rate Mode, 128kbps, Stereo).

Voice ratings
A sample of 80 participants (40 men and 40 women) aged 18 to 40 years ($M = 24.46$, $SD = 3.79$) listened to the vocal samples via circum-aural earphones (Superlux HD681F), set to constant amplitude, and judged them on age, attractiveness and femininity (in blocks, with voices randomized within each block). These presentations were created using Medialab software (Empirisoft Inc., New York, USA). All participants were recruited from the local student population in Göttingen (Germany) and reported to be German native speakers. Age estimations were provided using a free input field with accepted values set to a range of 1 to 99 years. Judgments of attractiveness and femininity were made on a 7-point scale (1 = not attractive/feminine to 7 = very attractive/feminine). Participants were instructed to judge the voices spontaneously and told that they would hear every voice only once in order to make their decision. All participants received a payment of 5€ for their participation. We used the mean ratings of age, attractiveness and femininity of all participants for the statistical analysis.

Facial visualizations
A further 80 participants (40 men and 40 women) aged 18 to 33 years ($M = 23.40$, $SD = 3.37$) listened to the vocal samples (randomized between participants) using the same technical setup as for the voice ratings, while watching a 1200 frames video clip (40 seconds at a rate of 30 frames per second) showing an extremely masculine looking male face (frame 0) that gradually changed into an extremely feminine looking female face (frame 1200) (Johnston, Hagel, Franklin, Fink, and Grammer, 2001). In this video, facial characteristics such as symmetry and hairstyle are kept constant, so that the predominant visual change is due to apparent masculinity/femininity. Participants were asked to indicate the face they felt each voice corresponded to, by moving the slider in the video time line forward and backward to that face. They were instructed to type in the face (frame) number using a free input field in Medialab. Voices were played in a loop until participants made a decision and clicked ‘continue’ to proceed. For the statistical analysis, we used the calculated means of frame numbers corresponding to each voice, as indicated by the participants.

Results
Table 1 reports descriptive statistics of female and male participants’ age, attractiveness, and femininity perceptions of women’s voices. There was a significant main effect of age group for all three attributes (age: $F = 487.21$; attractiveness: $F = 103.00$;
Facial visualizations of women’s voices

femininity: $F = 29.52$; all $p < .001$), with voices of young girls and adult women being perceived as significantly younger, more attractive and feminine than those of peri-/post-menopausal women. Omnibus post-hoc comparisons of age groups for the three attributes revealed significant effects throughout, with two exceptions: attractiveness and femininity perceptions of young girls’ and adult women’s voices did not differ significantly from one another (attractiveness: $p = .11$; femininity: $p = .99$).

Participants’ sex did not have a main effect on perception of women’s voices (age: $F = .88, p = .35$; attractiveness: $F = .47, p = .49$; femininity: $F = .69, p = .41$). There was a significant age group * sex interaction effect for attractiveness perception of women’s voices ($F = 3.78, p < .05$), but no significant effect for age and femininity (age: $F = 1.49, p = .23$; femininity: $F = .36, p = .63$). While men perceived the voices of young girls as relatively higher on the attractiveness scale, women did not. Female participants judged voices of adult and peri-/post-menopausal women higher on attractiveness than did male participants.

**Table 1.** Descriptive statistics ($M$ and $(SD)$) of female and male participants’ age, attractiveness and femininity perceptions of women’s voices.

<table>
<thead>
<tr>
<th>Age</th>
<th>Young Females</th>
<th>Young Males</th>
<th>Adult Females</th>
<th>Adult Males</th>
<th>Peri-/Post-menopausal Females</th>
<th>Peri-/Post-menopausal Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>18.52</td>
<td>19.00</td>
<td>25.81</td>
<td>24.47</td>
<td>47.19</td>
<td>44.55</td>
</tr>
<tr>
<td></td>
<td>(2.97)</td>
<td>(3.84)</td>
<td>(5.59)</td>
<td>(6.36)</td>
<td>(11.12)</td>
<td>(9.88)</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>3.77</td>
<td>4.01</td>
<td>4.16</td>
<td>4.01</td>
<td>2.78</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td>(0.89)</td>
<td>(0.92)</td>
<td>(0.74)</td>
<td>(0.80)</td>
<td>(0.92)</td>
<td>(0.85)</td>
</tr>
<tr>
<td>Femininity</td>
<td>4.34</td>
<td>4.60</td>
<td>4.52</td>
<td>4.56</td>
<td>3.53</td>
<td>3.58</td>
</tr>
<tr>
<td></td>
<td>(0.95)</td>
<td>(1.06)</td>
<td>(0.83)</td>
<td>(0.73)</td>
<td>(0.96)</td>
<td>(1.16)</td>
</tr>
</tbody>
</table>

Pearson correlations ($r$) of age, attractiveness and femininity perceptions of women’s voices showed significant correlations of attractiveness and femininity perceptions in males ($r = .46, p < .01$) and females ($r = .33, p < .05$), but no significant relationships of both attributes with age perception in either sex (males: age and attractiveness $r = .05, p = .75$; age and femininity $r = -.18, p = .26$; females: age and attractiveness $r = .14, p = .37$; age and femininity $r = -.07, p = .65$). Running these correlations for each of the three age groups separately revealed essentially the same results.

**Table 2.** Descriptive statistics ($M$ and $(SD)$) of female and male participants’ facial visualizations (in terms of frame numbers on the male to female video) of women’s voices.

<table>
<thead>
<tr>
<th>Frame Number</th>
<th>Young Females</th>
<th>Young Males</th>
<th>Adult Females</th>
<th>Adult Males</th>
<th>Peri-/Post-menopausal Females</th>
<th>Peri-/Post-menopausal Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>916.35</td>
<td>893.26</td>
<td>896.14</td>
<td>871.15</td>
<td>722.63</td>
<td>702.06</td>
<td>702.06</td>
</tr>
<tr>
<td>(133.08)</td>
<td>(136.24)</td>
<td>(132.25)</td>
<td>(123.07)</td>
<td>(142.41)</td>
<td>(168.52)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 reports descriptive statistics of male and female selections (in terms of frame numbers) of facial visualizations of women’s voices. There was a significant main effect of age group ($F = 110.07, p < .001$) with participants assigning higher frame numbers (i.e., higher facial femininity) to the voices of young women as compared to those of adult and peri-/post-menopausal women. There was no significant difference in facial visualizations of women’s voices from the young girls and adult women group ($p = .22$), but both groups differed significantly from that of the peri-/post-menopausal group (both $p < .001$); the latter received significantly lower frame numbers, indicating less facial femininity (see Figure 1).

![Figure 1.](image)

**Figure 1.** Facial visualizations of participants’ perceptions of women’s voices on the male to female movie ranging from A (frame 0) to B (frame 1200), and mean frame numbers (male and female responses together) of visualizations for the group of young girls (C, frame 904), adult women (D, frame 883) and peri-/post-menopausal women (E, frame 712).

Male and female participants did not differ significantly in their associations of faces on the continuum of masculine/feminine images corresponding to women’s voices ($F = .74, p = .39$). Also, we did not detect a significant interaction effect of age group * sex ($F = .01, p = .98$).

**Discussion**

One of the aims of the current study was to further understand people’s age, attractiveness and femininity perceptions of women of different age groups, i.e., young girls, adult women, and peri-/post-menopausal women, as most related studies investigating perception of women’s voices have studied college-age women. The groups in the current study were chosen as they represent characteristic ages through the female lifespan with regard to reproductive value (Buss and Schmitt, 1993). Overall, we found that the voices of young girls and adult women were perceived as being significantly younger, more attractive, and feminine than those of peri-/post-menopausal women. In comparison to female raters, male participants judged the voices of young girls to be more attractive, female raters perceived that adult voices were the most attractive. This may reflect differential selection pressures operating on female and male preferences, with males having evolved a stronger preference for youth than females. Although statistically significant, this effect was not very pronounced, and clearly requires further investigation in a larger sample than we had in the present study. We also acknowledge that the judgments could simply reflect the ages of the raters (aged 18-40) who were on average...
considerably younger than the women in the peri-/post-menopausal condition. These older voices might be rated more positively by an older group of raters, and future studies could address this issue.

Considering the relationships of people’s age, attractiveness, and femininity judgments of women’s voices, we found significant correlations in both male and female participants only for attractiveness and femininity, but neither of these attributes showed significant associations with age assessments. Thus, it seems that in evaluating vocal attractiveness of women, both females and males primarily rely on femininity perception. This result relates to previous findings reporting that vocal femininity predicts vocal attractiveness (Collins and Missing, 2003; Feinberg et al., 2008). Indeed, Feinberg et al. (2005) also found that ratings of vocal attractiveness were more related to perceptions of femininity than to perceptions of age. Nevertheless, the relationship of vocal femininity and attractiveness in women is thought to be moderated by age-related androgen and estrogen levels. Higher androgen levels are usually assumed with lower pitched voices (Abitbol et al., 1999; Feinberg et al., 2008) and there is evidence from systematic manipulations of voice pitch such that feminized versions of voices are perceived to be more attractive across societies (e.g., Apicella and Feinberg, 2009; Feinberg et al., 2008; Fraccaro et al., 2011; Jones et al., 2010; Puts et al., 2011). One limitation of this current study is that we did not measure circulating estrogen levels in our speaker sample, and so our suggestion that the changing preferences associated with age being dependent upon hormonal status remains speculative.

The second aim of our study was to investigate what people visualize when they hear unfamiliar voices. Previous research on the perception of women’s voices has mainly focused on associations between physical measures and/or facial photographs with certain vocal characteristics (e.g., Collins and Missing, 2003; Feinberg et al., 2005; Hughes, Harrison, and Gallup, 2002) or facial composites based on objective measurements of women’s voice pitch (Feinberg et al., 2005) and recorded people’s preferences. These studies have shown that women with attractive faces also tend to have attractive voices, and that vocal and facial femininity positively predicts vocal and facial attractiveness (Collins and Missing, 2003; Feinberg et al., 2005, 2008; Puts et al., 2011). Collectively, these findings lend support to the theory that men’s preferences for vocal and facial femininity are consistent across modalities because women’s faces and voices advertise common information about the senders’ hormonal status (Abitbol et al., 1999; Feinberg et al., 2005, 2008). Thus, it has been suggested that perceivers may use this cross-modal information in a way that may better inform their mate-choice decisions (see for review Feinberg, 2008).

Our results show that both men and women associate feminine looking faces with unfamiliar voices of young girls and adult women, but they tend to associate significantly less feminine looking faces with voices of peri-/post-menopausal women. As with age, attractiveness and femininity ratings of women’s voices, this effect was largely driven by the significant difference between facial visualizations of young girls’ and adult women’s voices and those of the peri-/post-menopausal group. In short, listening to the sound of young voices causes men and women to visualize feminine faces. We therefore suggest that there is an age-related cross-modality preference for femininity that applies to both female faces and voices, and is applicable to both sexes, given previous and our present results of
positive correlations of voice attractiveness and femininity. Our data did not show sex differences in female and male vocal judgments and facial visualizations, thus the potential consequences in terms of inter- and intra-sexual selection mechanisms remain to be investigated (see Puts et al., 2011). Indeed, O’Connor et al. (2011, 2012) show that women are more jealous of women with higher pitched voices, prefer their partners not spend time around women with higher pitched voices, and think that women with higher pitched voices are relatively more likely to cheat on their partners. Thus, prior work and the work presented here suggest that the relationship between women’s voice frequencies and mate quality is important for both male choice and female competition.

On a proximate level, there is evidence from neurobiology/neuroimaging studies for a link between face and voice processing, suggesting functional coupling between face and voice areas (Love et al., 2011; von Kriegstein and Giraud, 2006). For example, Smith et al. (2007) provided evidence for sensory integration of auditory and visual information in the representation of human gender, suggesting that auditory information influences face perception and voice pitch may have an additional adaptive role of improving face recognition and perception. In that study, androgynous faces paired with low pitch in the male range were perceived as males, whereas when such faces were paired with tones in the female pitch range they were judged as females. In addition, neuroimaging studies have found that the two specific brain areas for face- and voice recognition (i.e., the fusiform face area and the superior temporal sulcus) have direct structural connections and can exchange information, which is especially relevant in the context of person identification (Blank, Anwander, Kriegstein, 2011). From an evolutionary perspective, the ability to compare and match visual and auditory stimuli could indeed help to adapt the relative weight of individual sensory modalities to overall perception, and thus reduce mate choice errors (see also Møller and Pomiankowski, 1993).

In summary, our data on age, attractiveness and femininity perceptions of women’s voices add to those of previous studies by considering voices of three different age groups covering significant time periods in the female reproductive life span. We found that voices of peri-/post-menopausal women were perceived to be older, less attractive and less feminine than those of young girls and adult women, which may reflect a general adaptation towards the preference for youth. Our results support the hypothesis of matching information given by multiple signals, which is behaviourally relevant for optimizing human social perception.

Acknowledgements: This project was funded by the German Science Foundation (DFG), grant number FI 1450/6-1, awarded to B.F., as well as through the Institutional Strategy of the University of Göttingen. Two anonymous referees provided numerous valuable comments that helped improving an earlier version of the manuscript.

Received 28 November 2012; Revision submitted 30 January 2013; Accepted 8 February 2013
References


Facial visualizations of women’s voices


Facial visualizations of women’s voices


