



Sex matters: Neural correlates of voice gender perception



Jessica Junger^{a,b,*}, Katharina Pauly^{a,b}, Sabine Bröhr^a, Peter Birkholz^c, Christiane Neuschaefer-Rube^c, Christian Kohler^d, Frank Schneider^{a,b}, Birgit Derntl^{a,b}, Ute Habel^{a,b}

^a Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

^b JARA-Translational Brain Medicine, Jülich, Germany

^c Department of Phoniatrics, Pedaudiology and Communication Disorders, Medical School, RWTH Aachen University, Aachen, Germany

^d Department of Psychiatry, Neuropsychiatry Division, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

ARTICLE INFO

Article history:

Accepted 24 April 2013

Available online 6 May 2013

Keywords:

Gender differences

Voice perception

Opposite-sex effect

Pitch

fMRI

ABSTRACT

The basis for different neural activations in response to male and female voices as well as the question, whether men and women perceive male and female voices differently, has not been thoroughly investigated. Therefore, the aim of the present study was to examine the behavioral and neural correlates of gender-related voice perception in healthy male and female volunteers. fMRI data were collected while 39 participants (19 female) were asked to indicate the gender of 240 voice stimuli. These stimuli included recordings of 3-syllable nouns as well as the same recordings pitch-shifted in 2, 4 and 6 semitone steps in the direction of the other gender.

Data analysis revealed a) equal voice discrimination sensitivity in men and women but better performance in the categorization of opposite-sex stimuli at least in men, b) increased responses to increasing gender ambiguity in the mid cingulate cortex and bilateral inferior frontal gyri, and c) stronger activation in a fronto-temporal neural network in response to voices of the opposite sex.

Our results indicate a gender specific processing for male and female voices on a behavioral and neuronal level. We suggest that our results reflect higher sensitivity probably due to the evolutionary relevance of voice perception in mate selection.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Gender discrimination signals are evolutionarily important for the successful identification of adequate mates. Such signals are eclectic: ranging from visual signals, such as body (Tovee et al., 1999) and face symmetry (Little and Jones, 2006), to non-visual signals such as odor and vocal characteristics, especially pitch (Fitch, 2000; Jones et al., 2010), and seem to covariate with each other (Cornwell et al., 2004). Related sensory mechanisms are captured by signals relevant to fundamental reproductive needs. This seems to result in an adaptive allocation of attention to a counterpart of the opposite sex (Duncan et al., 2007). Correspondingly, there is widespread evidence, that humans prefer and react faster to signals of the opposite sex compared to those of the same sex in various modalities (Conway et al., 2008; Cornwell et al., 2004; Fischer et al., 2004; Hofmann et al., 2006; Jones et al., 2010; Proverbio et al., 2010; Spreckelmeyer et al., 2012). However, conflicting results have also been found reporting a same-sex advantage (Cellerino et al., 2004; Yamaguchi et al., 1995) or sex-independent effects (Latinus and Taylor, 2012; O'Toole et al., 1998).

The existence of a sex-dependent perceiver–stimulus interaction is therefore insufficiently resolved.

Among the variety of gender discrimination cues, one of the most important evolutionary signals in our environment is the human voice, which plays a key role in social interaction (Belin et al., 2004; Fitch, 2000) and mate selection (Apicella and Feinberg, 2009; Hodges-Simeon et al., 2010, 2011).

The widespread source–filter concept of human vocal production (Fant, 1960; Ghazanfar and Rendall, 2008; Titze, 2008) states that oscillations in the vocal folds located in the larynx (source) generate acoustic energy passing through the vocal tract (filter). There, (like a series of band-pass filters) formants or vocal tract resonances determined by time dynamics of length and shape of the vocal tract modify the emitted sounds and create the speech signal. Formants are relatively independent of the fundamental frequency (F0), which is determined by the vocal folds (Fitch, 2000). However, the speech signal does contain not only a linguistic message, but also multiple paralinguistic information (Honorof and Whalen, 2010). In the last years research focused more and more on these additional socially relevant voice properties independent of linguistic features of speech; especially emotion (McNally et al., 2001; Rodero, 2011), identity (Latinus and Belin, 2012; Latinus et al., 2011; Schweinberger et al., 2011), attractiveness (Bestelmeyer et al., 2012) and gender (Charest et al., 2013; Pernet and Belin, 2012; Zaska et al., 2009). Further, paralinguistic information enables voice to be an 'auditory face', referring to the similarity in processing complexity

* Corresponding author at: Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Pauwelsstr. 30, 52074 Aachen, Germany. Fax: +49 241 80 82401.

E-mail address: jbath@ukaachen.de (J. Junger).

and providing relevant information about social and biological characteristics across both modalities (Belin et al., 2011; Belizaire et al., 2007).

Normally, humans infer the gender of the speaker unconsciously (Simpson, 2009). This decision is mostly perceived from pitch (Jones et al., 2010), but there are much more phonetic and neurobiological differences related to male and female voices which are associated with the way in which they are produced from and perceived by different speakers (Simpson, 2009).

Acoustically, the primary difference between human male and female voices concerns fundamental frequency and formants (acoustic resonance) resulting from the anatomical sex differences in vocal folds and vocal tract emerging in puberty (Fitch and Giedd, 1999). Titze (1989) reported approximately 60% longer vocal folds in men resulting in a lower F0 as well as an about 15–20% longer vocal tract. A manipulation of source and filter properties reflected that, although F0 represents the most effective factor, both, F0 and formant frequency, are necessary to identify speakers' gender consistently (Hillenbrand and Clark, 2009). Thus, two extra-linguistic parameters are of major relevance for gender discrimination in human audition: the vocal cord dependent mean fundamental frequency (F0) and the formant frequencies depending on vocal tract dynamics (Poon and Ng, 2011; Sokhi et al., 2005; Titze, 1989; van Dommelen, 1990).

The average voice pitch varies about one octave (i.e. 12 semitones) between the sexes. The typical F0 for male speakers is ranging from 107 to 147 Hz, and for female speakers from 170 to 224 Hz (Dacakis, 2000; Simpson, 2009). In between is an overlapping range of 135–181 Hz, which can be associated with voices of both sexes. In this range the decision on a male or female voice also depends on other parameters like visual information or prosodic characteristics (Gelfer and Schofield, 2000; Oates and Dacakis, 1997). This “gender-ambiguous” range is centered on a “gender cut-off” F0 of around 160 Hz, with voices exceeding 160 Hz mostly being perceived as female and voices below 160 Hz being assigned to males (Oates and Dacakis, 1997; Spencer, 1988). Especially in this gender-ambiguous range formants get high importance as additional gender determining cues (Sokhi et al., 2005; Titze, 1989).

However, sex differences in vocalization are not unique to humans but have been reported in many other species (Hauser, 1996; Rendall et al., 2004). Differences involve sex-specific patterns of calling during mate selection as well as diverse related organic characteristics. For example, hormonal differences in songbirds can affect the motivation to sing and the structural development of specific brain nuclei regulating song production (Hauser, 1996). By means of single cell recording Narins and Capranica (1976) found greater cell numbers tuned to a specific mate call in female neotropical tree frogs than in males indicating a neural basis for auditory perception differences for mate calls between the sexes. This conforms to the finding that female frogs choose mates whose callings match the excitatory frequency of their auditory system best (Ryan et al., 1992). Thus, the sexually dimorphic characteristics of vocalization and its perception are important predictors of mating success and can give insight into the reproductive behavior of animals and humans.

Regarding cerebral differences between the sexes, there is widespread evidence of sex differences in human brain structure (Cahill, 2006; Wallentin, 2009) including the prefrontal cortex, the amygdala, the hippocampus and other temporal areas (Brun et al., 2009; Cahill, 2006), such as the primary auditory cortex (PAC; Rademacher et al., 2001). Further, an increasing number of functional neuroimaging studies report sex differences in neural activity patterns during different tasks, for example related to emotion (Kohn et al., 2011), audition (Brun et al., 2009), face perception (Derntl et al., 2010; Mather et al., 2010), stress (Goldstein et al., 2010) or the interaction of emotion and cognition (Koch et al., 2007) with or without performance differences (Cahill, 2006). In many cases, different processing strategies in both sexes result in comparable behavioral patterns.

Human voice as speech carrier is initially processed in the auditory cortex, namely BA41 and BA42 of the superior temporal gyrus including Heschl's gyrus, regions in the planum temporale along the right anterior superior temporal sulcus and middle temporal gyrus (Belin et al., 2000, 2002, 2004; Formisano et al., 2008; von Kriegstein et al., 2003, 2010). Sex differences in auditory perception are already observed in the precortical auditory pathway such as longer cochlear length in men resulting in longer cochlear travel times (Sato et al., 1991) or differential auditory sensitivity at different frequencies (Corso, 1963). Even in the brain stem women exhibit faster and increased responses than men in click-evoked auditory responses (Krizman et al., 2012). Similarly Ruytjens et al. (2007) found increased activation in the PAC in men as compared to women in reaction to music (vs. noise). However, the neural basis for sex differences in gender-specific voice processing is still unclear. To the best of our knowledge, there are only two studies dealing with the topic. Both found differential regional activation for male or female voice perception. Sokhi et al. (2005) studied a sample of 12 men and found greater activation in the right anterior superior temporal gyrus while listening to female voices compared to male ones. In contrast, male voices increased activation in the right precuneus. In contrast, Lattner et al. (2005) reported stronger neural responses to female as compared to male voices in the right supra-temporal plane, the right posterior superior temporal gyrus, left postcentral gyrus as well as bilateral inferior parietal lobe and insula in a mixed sample of 8 men and 8 women. Direct comparison revealed no sex differences in neural activation. However, both sample sizes were rather small preventing a thorough analysis of sex differences for female and male voices.

Therefore we investigated a larger sample of healthy women and men and systematically varied fundamental frequency and formants of originally male and female voices to explore the neural correlates of gender-related voice perception. Based on previous results (Lattner et al., 2005; Sokhi et al., 2005) we expected a differential neural processing of original male and female voices in voice-selective regions of the brain, such as the superior temporal gyrus and middle frontal cortex, interacting with the sex of listener, namely increased activation in response to voices of the opposite sex due to increased attention to evolutionary relevant stimuli. Moreover and based on previous

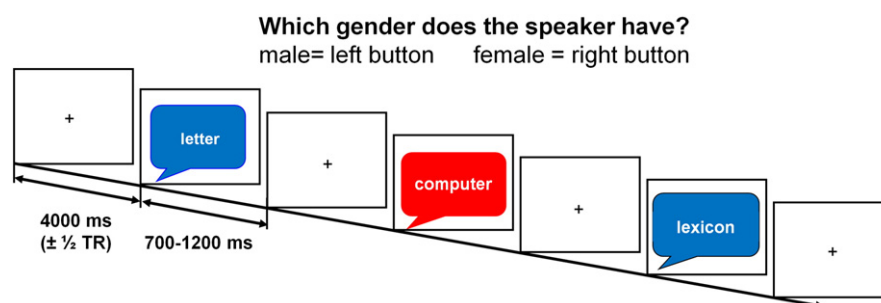


Fig. 1. Order of events for the experimental paradigm with an illustration of acoustic presented male (blue) and female (red) voices.

literature (Charest et al., 2013; Schweinberger et al., 2008) we hypothesized an increased recruitment of additional frontal and cingulate regions in response to gender ambiguous voices (due to increased cognitive demands and monitoring processes). Therefore, we implemented a task presenting different voice morphing levels (see also: Belizaire et al., 2007; Mullenix et al., 1995). Starting from unambiguous male and female voices we increased ambiguity by morphing voices stepwise (2, 4 and 6 semitones) into the direction of the opposite sex. This modification of the degree of difficulty of gender recognition might represent an especially sensitive tool for the investigation of sex differences in voice perception. In this way we intended to not only investigate differences in gender-specific voice processing between men and women, but also to show basal responses of the brain to evolutionary determined cues in human voice.

Methods

Subjects

In total, 41 healthy subjects were examined. Two subjects were excluded subsequently due to movement artifacts. Hence, data from 39 participants (19 females, 20 males) were included in the final analyses. On average, men and women did not differ significantly regarding age (men: 32.35 years, SD = 10.26; women: 33.16 years, SD = 12.34, $t(37) = -0.223$, $p = 0.825$), education (men: 15.00 years, SD = 2.92; women: 14.95 years, SD = 3.21, $t(37) = 0.054$, $p = 0.958$) or crystallized verbal intelligence quotient (men: 112.45, SD = 14.71; women: 112.21, SD = 16.14, $t(37) = 0.048$, $p = 0.962$). Aside from one left-handed participant in each group all subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971).

Subjects were screened by means of the Structured Clinical Interview for DSM-IV (Wittchen et al., 1997) to ensure the exclusion of mental disorders of axis I. Subjects with neurological disorders and other medical conditions which could affect the cerebral metabolism as well as with first degree relatives with a history of mental or neurological illness were also excluded.

The study was approved by the local Institutional Review Board of the Medical Faculty of the RWTH Aachen University. All participants gave their written informed consent and were financially reimbursed for their participation.

Stimuli

The stimuli consisted of 10 male and 10 female voices. Six German target words were recorded – three nouns spoken by half of each group, respectively resulting in 30 natural male and 30 natural female stimuli. In other words, the six stimulus words as well as the male and female speakers were divided in halves, i.e. three of the six words were spoken by five male and the other three by five female speakers.

In addition, as described in detail below, three modified (morphed) versions of each natural stimulus were created to make the natural stimuli progressively more gender-ambiguous. Hence, we got 240 stimuli in total.

The six target words were trisyllabic, emotionally neutral nouns selected from a bibliographic database (Baayen et al., 1995). Trisyllabic words were selected because they were found to be adequate for voice gender discrimination with an error rate of less than 1% (Lehmann et al., 2010). For a natural and consistent pronunciation, the target words were spoken in the context of the carrier sentence "I said ..." and then cut out. The utterances were digitally recorded with a high-quality microphone (type C 2000 B by AKG) at a sampling rate of 44100 Hz and a quantization of 16 bit in a sound proofed room using the software tool Audacity (<http://audacity.sourceforge.net>). After cutting out the stimuli from the carrier sentences, their amplitudes were normalized. To avoid any discontinuities at the beginning and the

end of the cut-out stimuli, their onsets and offsets were smoothed by multiplying them with the left/right half of a 30 ms long Hann window. From multiple repetitions of the recordings, instances of the stimuli were selected such as to match in duration across conditions and sex (mean duration female voices: 946.43 ms (SD = 159.885), male voices: 996.66 ms (SD = 112.903); $t = -1.445$; $p = 0.153$). The average fundamental frequency of the natural stimuli was 207.09 ± 32.11 Hz for female speakers and 128.77 ± 23.42 Hz for male speakers.

For each of the natural stimuli, three modified versions were produced to make the original male stimuli sound progressively more female and the original female stimuli sound progressively more male. The modifications were done using the "change gender" function of the software Praat (Version 5.2.03 www.praat.org; Boersma and Weenink, 2010). As stated before, adult women are characterized by a mean pitch that is about one octave (= 12 semitones) above that of men, and a vocal tract that is proportionally shorter than that of men (Fant, 1960; Fitch and Giedd, 1999). Therefore, each natural stimulus was modified to approach the gender-ambiguous voice region in three steps by shifting the pitch contour and changing the formant structure that reflects the vocal tract length. Since the average frequency of a formant and vocal tract length are roughly inversely related, the formant shift ratio was set to the inverse of the vocal tract length factors given above. Original male stimuli were modified by shifting the pitch contour up by 2, 4, and 6 semitones (st), and at the same time reducing the vocal tract length by the factors 0.97, 0.94, and 0.91, respectively. Accordingly, original female stimuli were modified by shifting the pitch contour down by 2 st, 4 st, and 6 st, and at the same time increasing the vocal tract length by the factors 1.03, 1.06, and 1.09, respectively.

All natural and modified stimuli were rated in a pre-study by an independent sample of 15 male and 15 female subjects in order to validate the stimuli used in the fMRI experiment regarding naturalness on a 10 point scale with 10 being most natural (male voices: mean (SD): 0 st: 6.27 (1.63), 2 st: 5.25 (1.15), 4 st: 4.68 (1.22), 6 st: 4.39 (1.18); female voices: 0 st 7.09 (1.90), 2 st 6.25 (1.14), 4 st 5.83 (1.10), 6 st 5.19 (1.18)). By means of variance analyses, no significant differences were detected between the original voices and their morphed equivalents (female: all $p > 0.245$; male: all $p > 0.178$). In addition, there were no rating differences between men and women for any condition (all: $p > 0.151$).

Task

Stimuli were presented via an MRI compatible sound system by means of electrostatic headphones with Presentation 14.2 software (<http://www.neurobs.com>). Loudness was adapted individually (ca. 70 dB on average) to assure comfort, intelligibility and audibility given the background noise of the scanner.

Every condition appeared ten times in mini-blocks consisting each of 3 nouns spoken by 3 different speakers within a hybrid design resulting in 80 blocks. The order was pseudo-randomized and balanced so that no speaker and no word was heard consecutively. The first block was preceded by a short instruction. Within a block an inter-stimulus interval (ISI) of 4 s (plus randomized jitter of maximal $\pm 1/2$ TR) was used. Between blocks there was an ISI from 11 s to 16 s (\pm randomized jitter of maximal $1/2$ TR). Therefore the total length of the paradigm was 27.2 min.

The task was to determine the gender of each stimulus by pressing one of two target buttons of an MR-compatible response pad with the right index (male voice) or middle finger (female voice) as fast as possible (Fig. 1).

Analysis of the behavioral data

All behavioral data were analyzed using SPSS 18.0.0 (SPSS Inc., Chicago, IL). Since the behavioral raw data were not normally

distributed, nonparametric Friedman rank tests (Friedman, 1937; Mann and Whitney, 1947) were performed to detect mean rank differences between conditions in response accuracy and reaction times (RT). Post-hoc Mann–Whitney–U tests for between-group comparisons and Wilcoxon tests for within-group comparisons for each sex group separately as well as for the whole group were calculated to decompose significant effects (Mann and Whitney, 1947; Wilcoxon, 1945).

Task performance was assessed by the amount of hits (true positive answers) and false negative responses (due to the characteristics of the two-alternative discrimination task the inclusion of true negative and false positive responses would not result in additional information given that indicating a “female” voice in response to a male stimulus can be interpreted as a false negative answer during the male condition or as a false positive reaction in the female condition).

To further explore performance based on Signal Detection Theory, discrimination sensitivity and potential response biases (D’Ausilio et al., 2011) for each morphing level were assessed by means of d' and log beta ratios ($\log \beta$) for two-alternative discrimination tasks (Wickens, 2002). The latter gets zero at the point exactly between both distributions, positive when biased to the first answer alternative (here: “male voice”) and negative when biased to the other (“female voice”). One-sample t-tests were calculated for each sex group separately against H_0 with a test value = 0 to see if discriminability differed from chance level and if there is a significant position bias. Furthermore, two-sample t-tests were computed to compare the groups.

All post-hoc tests of the behavioral data were Bonferroni corrected for multiple comparisons.

fMRI data acquisition

Functional images were acquired on a 3T Siemens Trio MR Scanner (Siemens Medical Systems, Erlangen, Germany) at the Department of Psychiatry, Psychotherapy and Psychosomatics of the University Hospital of the RWTH Aachen University using echo-planar imaging (EPI) sensitive to BOLD contrast (T_2^* , voxel size: $3.1 \times 3.1 \times 3.1 \text{ mm}^3$, distance factor 15%, GAP 0.5 mm, 64×64 matrix, FoV: $200 \times 200 \text{ mm}^2$, TR = 2 s, TE = 30 ms, $\alpha = 76^\circ$). 36 slices covered the entire brain. Transversal slices were acquired in ascending order. Image acquisition was preceded by 6 dummy images in order to account for magnetic field saturation effects. They were discarded before preprocessing resulting in 785 volumes for each subject.

fMRI data analysis

Image analyses were performed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>) implemented in MATLAB 2010b (Mathworks, Sherborn, MA). In a first step, the images were corrected for head movement by realigning them to the first image. Afterwards, the realigned images were spatially normalized into the anatomical MNI space by coregistering them to the SPM8 tissue probability map template for grey matter and applying the non-linear unified segmentation approach (Ashburner and Friston, 2005). Normalized images were smoothed with an 8 mm full-width-at-half-maximum (FWHM) Gaussian kernel to compensate for inter-subject variations in brain anatomy. Low frequency noise was filtered out using a high pass filter of 128 Hz.

Afterwards, on an individual level, all 8 experimental conditions were separately modeled in an event-related design relying on the canonical hemodynamic response function. First-level contrasts were taken to a second-level and entered in a mixed-effects general linear model (GLM) using a flexible factorial design implemented in SPM8 with subjects as random effect (between group factor: sex) and conditions as fixed effects (within-subject factors: sex and morphing level of the voice (i.e. unmorphed original voice, morphed

by 2, 4 or 6 st)). Further, movement parameter regressors and the individual mean reaction times were included as nuisance covariates.

Moreover, to assess the effects of the parametric modulation of the stimulus morphing, contrasts for male and female voices were weighted linearly ascending ($0\text{st} * -3 < 2\text{st} * -1 < 4\text{st} * 1 < 6\text{st} * 3$) and mean centered according to their morphing level.

Based on this model, separate contrasts were computed. First, two contrasts for the interaction of group/sex of the listener and sex of the (original) voice assessing the hypothesized opposite sex effect, namely: 1) [men 0w > men 0m] > [women 0w > women 0m], 2) [women 0m > women 0w] > [men 0m > men 0w] (note: due to the subtraction method in the construction of contrasts, the first interaction is arithmetically equivalent to [women 0m > women 0w] > [men 0m > men 0w] and the second to [men 0w > men 0m] > [women 0w > women 0m]). The underlying effects were decomposed by parameter estimates. Mean beta values and standard errors of each peak voxel were extracted for each of both voice conditions and in each group separately in order to segment significant results by means of post hoc two-sample t-tests comparing all voice and sex sub-conditions with each other using GraphPad (www.graphpad.com).

Second, effects of the morphing degree were assessed across the whole group and compared between the sexes.

Finally, a correlation analysis reflected the relationship between the amount of errors and the activation alteration due to an increased gender ambiguity of the voices (i.e. the morphing degree).

A voxel-level threshold of $p < 0.05$ FWE corrected for multiple comparisons (extent threshold: 20 voxels) was adopted for the whole group and the correlation analyses. A more liberal Monte Carlo corrected threshold was used for the analysis of sex group differences, since interaction effects are more subtle and represent the difference of a difference with less statistical power. Monte Carlo simulations were computed using AlphaSim by Ward (2000) implemented in AFNI 2011 (Cox, 2012). Assuming a per voxel probability threshold of $p = 0.001$, after 1.000 simulations a cluster size of 19 contiguous resampled voxels was indicated to correct for multiple comparisons at $p < 0.05$.

Results

Behavioral results

Hits and errors

The Friedman test revealed significant differences regarding the percentage of hits ($\chi^2(7) = 198.37$, $p = 0.001$) and errors ($\chi^2(7) = 199.35$, $p < 0.001$) across all morphing and sex of the voice conditions. No general sex ($p = 0.888$) and voice gender ($p = 0.625$) effects were found. In contrast, stimulus morphing decreased percentage of hits in the whole group (as well as in both sex groups separately (all $p < 0.001$)).

Further looking at both sexes separately and in contrast to whole group analysis, an effect of sex of the voice was found with more correct answers for voices of the opposite sex in men ($z = -2.465$, $p < 0.014$) and women ($z = -3.018$, $p < 0.003$), respectively.

Directly comparing both groups, significant sex differences were found for male and female voices (both $z > -3.584$, $p < 0.001$). Both groups showed better performance for voices of the opposite sex. However, no group differences were found concerning the morphing levels (all $p > 0.224$). Further, post hoc comparisons of each task sub-condition reflected significance for all differences between men and women (at least on a statistical threshold of $p < .05$), although only sex differences for the 4st and 6st conditions survived Bonferroni correction (see Table 1, Fig. 2).

Given the two-alternative discrimination task, the opposite patterns were found for error rates (accordingly providing no additional information).

Table 1

Mean percentage of hits, errors, reaction times for correct responses in seconds, discrimination sensitivity (d' -prime) and answering bias ($\log \beta$) in response to male and female voices of the different morphing steps in semitones (st) for women and men.

	Men	Women		Men	Women	
	Mean (SD)	Mean (SD)	p	Mean (SD)	Mean (SD)	p
Hits (%)						
<i>Male voice</i>				<i>Female voice</i>		
0st	97.00 (3.73)	99.29 (1.78)	0.020*	97.66 (2.88)	94.21 (6.29)	0.032*
2st	90.66 (7.48)	96.84 (4.07)	0.003*	91.16 (10.83)	80.35 (11.75)	0.002*
4st	63.00 (15.02)	83.15 (12.78)	<0.001**	81.66 (9.64)	65.78 (15.98)	0.001**
6st	36.50 (17.18)	62.63 (15.65)	<0.001**	58.66 (16.27)	36.49 (15.93)	<0.001**
Error rate (%)						
<i>Male voice</i>				<i>Female voice</i>		
0st	2.66 (3.17)	0.52 (1.67)	0.009*	1.83 (2.75)	5.78 (6.26)	0.006*
2st	9.00 (7.26)	3.15 (4.07)	0.004*	8.33 (10.84)	19.29 (11.78)	0.002*
4st	36.50 (16.87)	16.84 (12.78)	<0.001**	18.00 (9.57)	33.68 (15.98)	0.001**
6st	63.16 (16.87)	37.19 (15.72)	<0.001**	41.00 (16.36)	63.15 (15.88)	<0.001**
Reaction time						
<i>Male voice</i>				<i>Female voice</i>		
0st	1.22 (0.19)	1.06 (0.15)	0.011*	1.25 (0.23)	1.07 (0.16)	0.011*
2st	1.32 (0.29)	1.14 (0.14)	0.009*	1.32 (0.27)	1.18 (0.24)	0.049*
4st	1.51 (0.29)	1.22 (0.13)	0.001**	1.34 (0.25)	1.27 (0.21)	0.448
6st	1.65 (0.37)	1.35 (0.16)	0.008*	1.47 (0.29)	1.38 (0.27)	0.206
D-prime				Log β		
0st	3.87 (0.34)	3.78 (0.41)	0.495	-0.15 (0.73)	0.63 (0.75)	0.002*
2st	3.03 (0.48)	2.84 (0.40)	0.179	-0.18 (1.32)	1.22 (1.08)	0.001**
4st	1.35 (0.32)	1.53 (0.43)	0.128	-0.08 (0.41)	0.37 (0.49)	0.003*
6st	-0.15 (0.45)	-0.03 (0.29)	0.339	-0.08 (0.41)	0.37 (0.49)	0.003*

* $p < 0.05$.

** $p < 0.002$ (Bonferroni corrected).

Discrimination and response bias

Two-sample t-tests on d' showed no significant difference between men and women in discrimination performance for all 4 morphing degrees (see Fig. 2 and Table 1). Further, one-sample t-tests revealed that discriminative ability was above chance level for all conditions ($p < 0.001$) except the highest morphing level (6st – with $p = 0.147$ for men and $p = 0.608$ for women). In general, discrimination sensitivity decreased with increasing morphing degree (see also Table 1).

Two-sample t-tests on $\log \beta$ indicated sex differences in all 4 morphing degrees (see Table 1). However, only group differences for voices morphed by 2st also survived Bonferroni correction of (of $p < 0.002$) (see also Table 1). While men revealed no significant position bias (all $p \geq 0.370$) women were biased to indicate male voices (all $p \leq 0.004$).

Reaction times

In addition, Friedman tests revealed significant differences in reaction time (RT) across all morphing and sex of the voice conditions ($\chi^2(7) = 172.94$, $p = 0.001$). For the whole group, a general sex effect was found with women responding overall faster than men ($z = -2.529$, $p = 0.011$), which was also true for all morphing levels separately (all $p < 0.022$). Further, an increase in stimulus morphing decreased reaction times (all $p < 0.005$) and overall participants reacted faster to female than to male voices ($z = -1.996$, $p < 0.046$).

Looking at both sexes separately, an increase in RT was found with increased stimulus morphing with significant differences between all morphing levels ($p < 0.002$) except 4st compared to 6st (men: $z = -2.053$, $p = 0.040$; women: $z = -2.495$, $p = 0.013$), where differences did not survive Bonferroni correction. Sex of the voice had a significant effect in men ($z > -2.688$, $p = 0.007$), but not in women ($z = -0.241$, $p = 0.809$) with faster responses to female than to male voices in male participants.

Comparing both groups directly, sex groups differed significantly regarding male ($z = -2.894$, $p = 0.003$), but only by trend in response to female ($z = -1.939$, $p = 0.054$) voices. While women responded faster than men in both conditions men responded faster to female than to male voices. In the direct comparison of each task sub-condition women responded significantly faster than men in all conditions, except for the 4st and 6st female voice condition at least on a statistical threshold of $p < .05$ (Table 1). However, only group differences for male voices morphed by 4st survived the conservative Bonferroni correction.

To take these significant sex differences into account, RTs were included as covariate of no interest into the brain image analyses.

Functional imaging results

Effects of sex and original voice gender

Interaction analyses showed stronger activation in men compared to women for the processing of female vs. male original voices in the right superior prefrontal gyrus similar to medial prefrontal cortex (MPFC), left medial orbitofrontal cortex (OFC) and the left middle temporal gyrus (MTG, s. Table 2) extending to angular gyrus.

In the next step, the extracted mean beta values and standard errors from each of these three regions were compared with each other for each group and each original voice condition by means of post hoc t-tests in order to decompose the interaction effect.

Post hoc comparisons in the MPFC revealed stronger activation during the processing of female voices in men as compared to women ($t(37) = 3.072$, $p < 0.004$). Stronger activation was also shown for female voices compared to male voices in men ($t(38) = 2.166$, $p < 0.037$) and (by trend) for male voices compared to female voices in women ($t(38) = 1.987$, $p < 0.055$).

Comparisons in the OFC showed less deactivation in response to female voices in men as compared to women ($t(37) = 2.1525$,

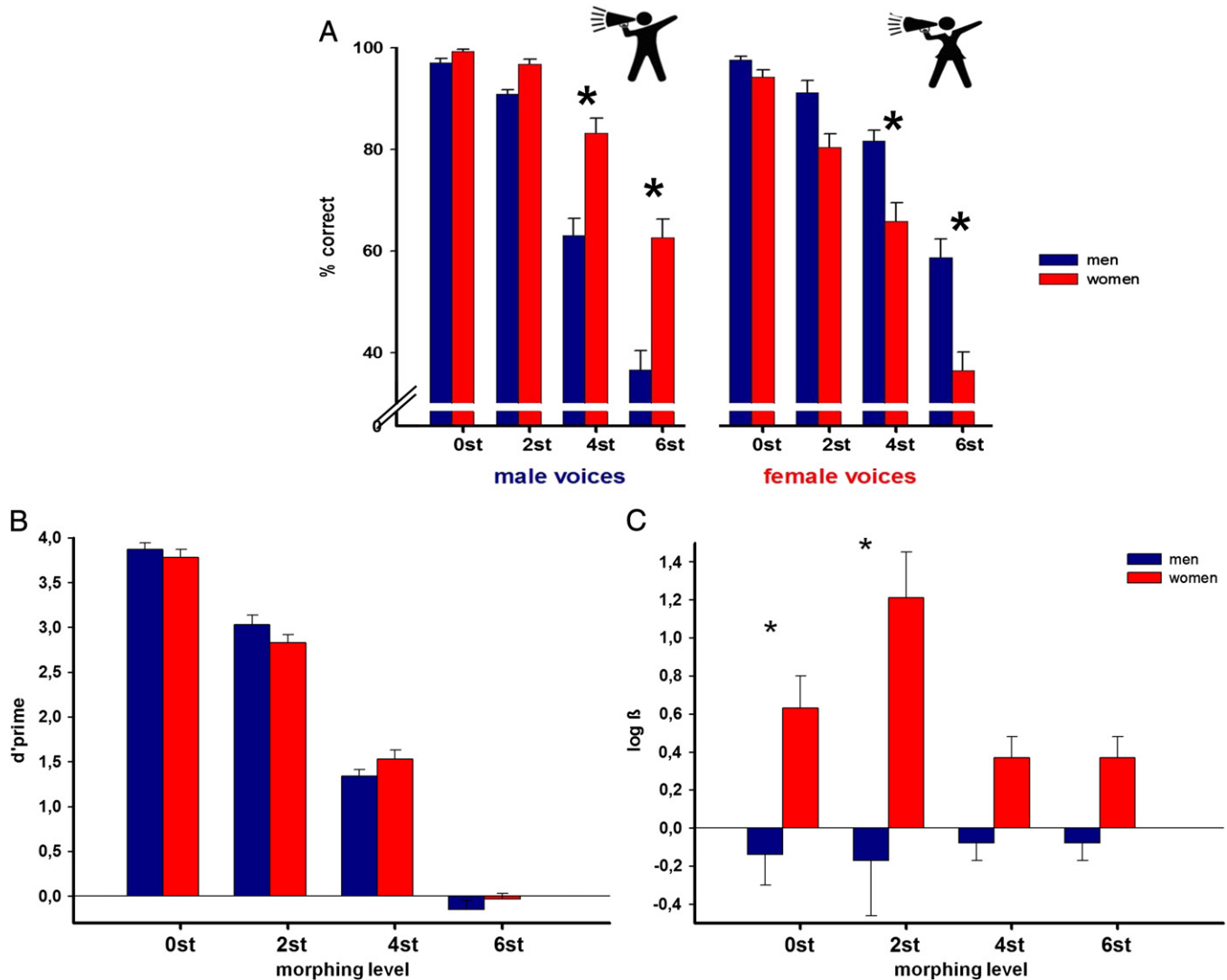


Fig. 2. A. Performance (% correct with standard error bars) in response to male (left) and female (right) voices of the different morphing steps in semitones (st) in men (blue) and women (red). B. Gender discrimination sensitivity (d' -prime with standard error bars) and C. response bias ($\log \beta$ with standard error bars) for each morphing step in men (blue) and women (red) with positive values representing bias to choose male voices and negative values representing bias to choose female voices. Significant differences are marked by asterisks.

$p < 0.038$). In addition, less deactivation was found for male voices as compared to female voices in women ($t(38) = 2.402$, $p < 0.022$).

Post-hoc comparisons in the MTG/angular gyrus, finally, exhibited less deactivation during the processing of female voices in men as compared to women ($t(37) = 2.987$, $p < 0.005$). Furthermore, less deactivation was revealed for male as compared to female voices in women ($t(38) = 2.611$, $p < 0.013$) (Fig. 3).

In all three regions there was no significantly stronger activation in women compared to men, in response to male compared to female voices in men, or in response to female compared to male voices in

Table 2

Stronger activation/less deactivation in men compared to women for the processing of female vs. male original voices with no significant results for the opposite interaction ($[\text{women } 0m > \text{women } 0w] > [\text{men } 0m > \text{men } 0w]$; MNI coordinates, $p < 0.05$ Monte Carlo corrected, $k = \text{cluster extension}$).

Brain region	L/R	x	y	z	k	t
Medial pre frontal cortex	R	15	56	16	120	4.48
Middle temporal gyrus (extending to angular gyrus)	L	-54	-76	13	24	4.19
Medial orbitofrontal cortex	L	-12	35	-14	27	4.06

women. Note that only activation in the MPFC and the MTG/angular gyrus during the processing of female voices in men as compared to women survived Bonferroni correction ($p < 0.008$).

Effects of voice gender morphing

The parametric weighting of the linearly increasing morphing degree revealed activation in the inferior frontal gyri bilaterally extending to the insula as well as in the mid cingulate cortex (MCC) for the whole group (Table 3 and Fig. 4).

Men showed stronger activation in right superior and middle frontal gyri than women for increasing stimulus morphing (Table 3). Parameter estimates revealed that effects resulted from an increase in activation with increasing morphing degree in men not reflected in women (see Fig. 5). No stronger activation was found in women contrasted to men.

Correlation analyses exhibited a negative association between errors and activation increase with increasing morphing degree in the right insula, the right middle temporal gyrus and brain stem as well as left middle temporal pole, left hippocampus, and left paracentral lobe in women. However, no correlations were found in men (see Table 4).

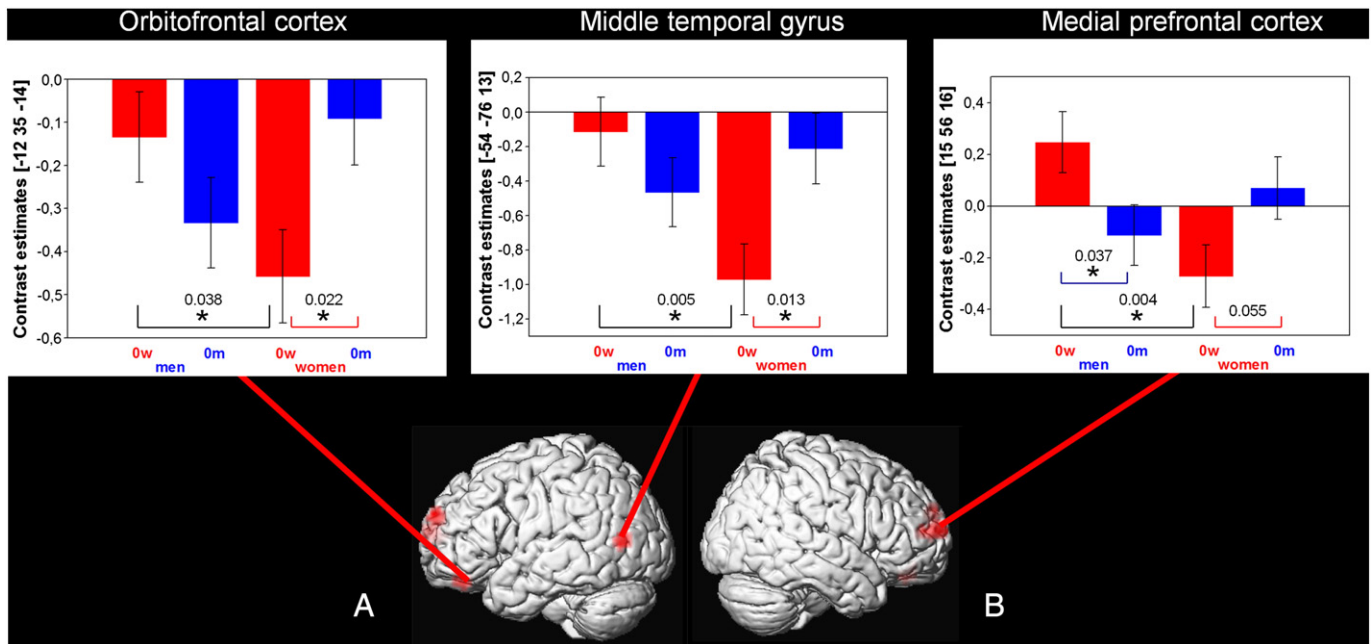


Fig. 3. Interaction between gender and original voice sex ($p < 0.05$ Monte Carlo corrected). Parameter estimates are shown separately for male (0m) and female (0w) voices for men and women; significant differences are marked by asterisks ($*p < 0.05$, $**p < 0.008$ Bonferroni corrected). A: left hemisphere, B: right hemisphere.

Discussion

The processing of voices is a basic cognitive and phylogenetic ability (Belin et al., 2004; Lattner and Friederici, 2003). Studying sex differences in its cerebral substrates might provide further insight not only into sex-specific voice perception, but also into the processing of gender-related information in general. For this purpose, the aim of the present study was to investigate the neural correlates of voice perception in men and women in dependence of gender-specific voice parameters. We could show better performance in the categorization of and stronger activation in response to opposite-sex stimuli in men compared to women as well as a response bias towards male voices together with stronger activation in response to male as compared to female voices in women. Moreover, men additionally recruited prefrontal areas with increasing morphing degree compared to women. Women, on the other hand, made more classification errors with decreasing activation in middle temporal and insula regions.

Biological relevance of opposite-sex voices

Opposite-sex effect on performance

We used the percentage of hits and errors, reaction times, discrimination sensitivity (d') and position bias ($\log \beta$) as behavioral measures of voice gender perception performance. First, when

comparing the amount of hits (and errors) in men and women, independent of the morphing level men revealed significantly better performance in response to female voices and vice versa. This effect was also reflected in the within-sex effects concerning the sex of the voice, i.e. both groups categorized voices of the opposite sex more accurately than the voice of the own sex. Second, women reflected an opposite-sex response bias, tending to more often choose male sex in their answering pattern, which partly may have resulted in the increase in true positive answers for male stimuli (given the two alternative forced-choice task) and therefore might have influenced the opposite sex performance effect in women. Third, men and women did not differ in their voice sex discrimination sensitivity (d').

Fourth, women generally showed faster responses than men. Looking at both groups separately, men reacted significantly faster in response to female as compared to male voices. This annulled group differences for female voices. According to that, men seem to benefit from the opposite-sex effect behaviorally as reflected by shorter reaction times for female as compared to male voices. Shorter reaction times in women, on the other hand, might be also related to the response bias, which might have resulted in faster reactions.

The fact that at least men revealed an opposite-sex effect in performance is in line with findings from previous studies in different domains revealing that men prefer women's pheromones (Cornwell et al., 2004) and faces (Feinberg, 2008) to male ones. Moreover, Hofmann et al. (2006) showed that men entitle the facial attractiveness of female faces faster than women and vice versa. According to that, they found an opposite-sex performance effect which we could replicate for voices.

Interestingly, and in contrast to Brown and Perrett (1993), Yamaguchi et al. (1995) reported an own-sex effect in face perception in an Asian sample. Authors explained conflicting results with a possible incomparability between Asian and Caucasian samples and underlined the need for further research on this topic. However, also Cellerino et al. (2004) found that women recognized female faces better than men – at least at certain pixelation levels. Although authors did not report on discrimination sensitivity or response bias measures, this raises the question of evolutionary differences regarding the processing of visual and auditory material and the influence of different social stimuli.

Table 3

Activation peaks (MNI coordinates) and cluster extension (k) for a linear increase of voice morphing regarding gender identity in the whole group ($p < .05$ FWE, cluster size > 20 voxels) and for men contrasted to women ($p < .05$ Monte Carlo corrected).

Contrast	Brain region	L/R	x	y	z	k	t	
Increasing degree of morphing	Overall	Inferior frontal gyrus (triangularis)	R	30	29	-5	458	10.28
		Insula	L	-30	26	-5	287	9.44
	Middle cingulate cortex	R	9	20	43	408	8.50	
	Inferior frontal gyrus (triangularis)	L	-48	20	25	69	5.57	
Men > women	Superior frontal gyrus	R	15	5	52	45	4.59	
	Middle frontal gyrus	R	30	29	31	61	3.82	

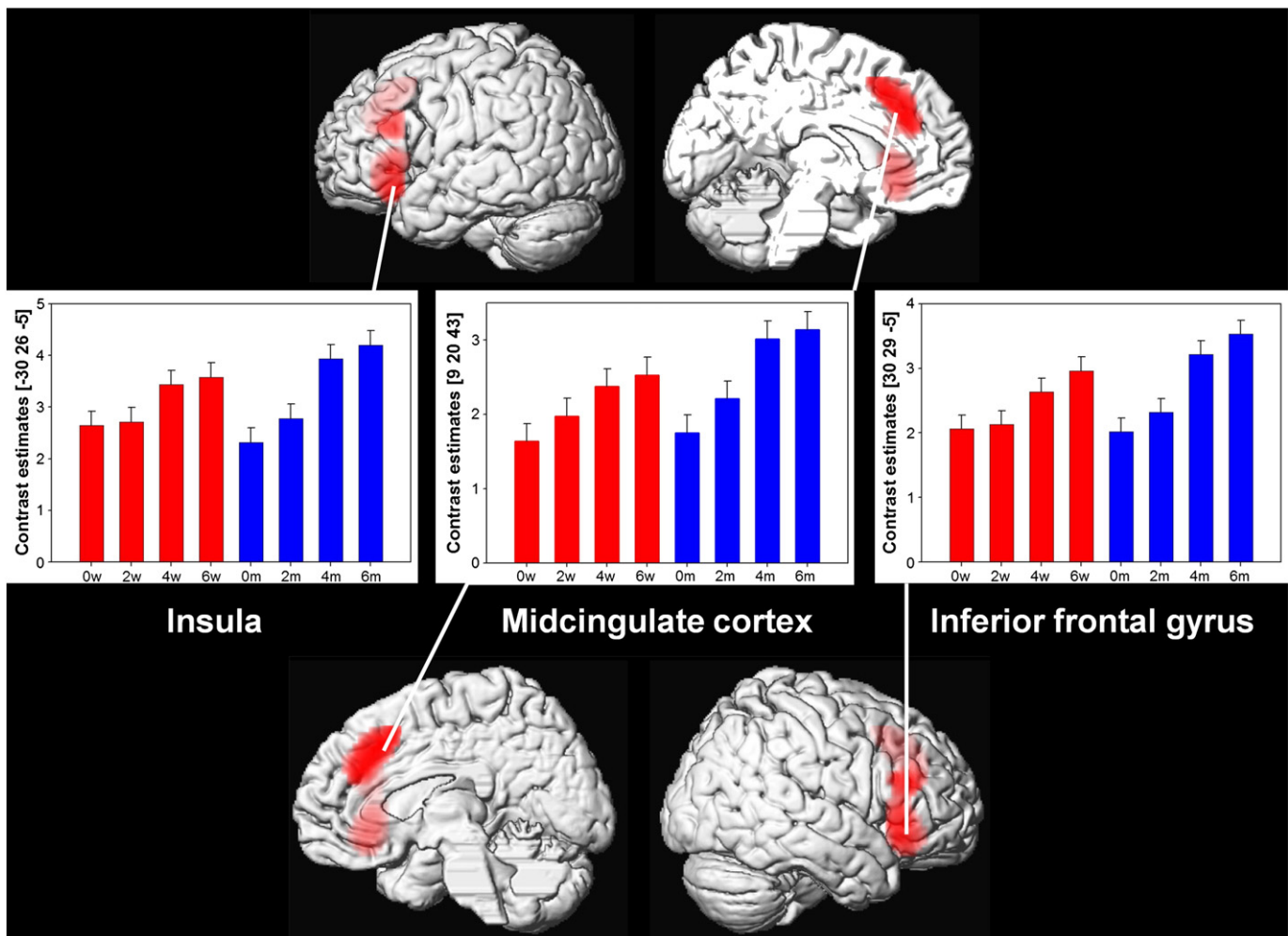


Fig. 4. Activation pattern for a linear increase of voice morphing regarding gender identity in the whole group ($p < 0.05$ FWE corrected, extent threshold = 20 voxels).

Nonetheless, an opposite-sex performance effect for voices in general (Feinberg et al., 2008) and especially for voice pitch (Jones et al., 2008, 2010) supports our findings. In addition, there is evidence for an association between men's and women's perceived vocal attractiveness and putative indices of fertility and health (Hughes et al., 2008).

Not all groups found an interaction between the listener's sex and the sex of the voice (Bruckert et al., 2010; Latinus and Taylor, 2012). However, subtle differences may only become apparent in larger sample sizes with more statistical power (both, Bruckert et al. (2010) and Latinus and Taylor (2012) investigated samples with not more than 13 female and 12 male participants).

In sum, our behavioral findings may be interpreted in terms of an evolutionary perspective suggesting that individuals attend more strongly to potential mates (Feinberg, 2008). This perspective of significance due to evolutionary reasons, possibly more pronounced in men, is in agreement with neurobiological and behavioral studies reporting that the perceived physical attractiveness can be modulated by the individual social interest in the evaluated person (Jones et al., 2008; O'Doherty et al., 2003). According to this interpretation, the social significance of a potential mate may be higher and result in an unconscious preference and better attendance to sexually differentiating cues, such as a speaker's voice, which may have led to our results of better recognition of voices of the opposite sex.

Opposite-sex effect on prefrontal activation

In line with the idea of an opposite-sex effect, our data resemble evidence from other modalities demonstrating increased activation

in men contrasted to women in the OFC in response to erotic pictures (Sabatinelli et al., 2004) and other emotionally relevant stimuli (Aleman and Swart, 2008), as well as in the MPFC in emotional (Shirao et al., 2005) and social decision tasks (Krach et al., 2009). There is a long line of evidence that prefrontal brain regions are involved in social judgment and attention. Further, both OFC and MPFC are closely linked to limbic structures critical for affective and motivational processes (Miller and Cohen, 2001) and recruited during cognitive reappraisal to regulate emotional responses (Ray and Zald, 2012).

Thereby, especially OFC is involved in the evaluation of emotional relevant stimuli across various sensory modalities (Hornak et al., 2003) and plays a key role in affective appraisal by reflecting the attribution of relative values to emotionally relevant concurrent voices (Sander et al., 2005). Further, there are studies showing especially responses in the OFC to faces of the opposite compared to faces of the same sex in heterosexual (Davey et al., 2010; O'Doherty et al., 2003) as well as to faces of the same sex in homosexual participants (Ishai, 2007). These findings support the notion of voice as homolog to faces and provide further evidence for an auditory face model of voice perception (Belin et al., 2011).

According to the model of emotion regulation by Phillips et al. (2008) the OFC is involved in the initial processing and identification of emotional stimuli and thereby closely interconnected with the MPFC. Inter alia, the latter is suggested to be involved in the computation of expected reward and cognitive aspects of emotion processing, such as attention to or identification of emotional material

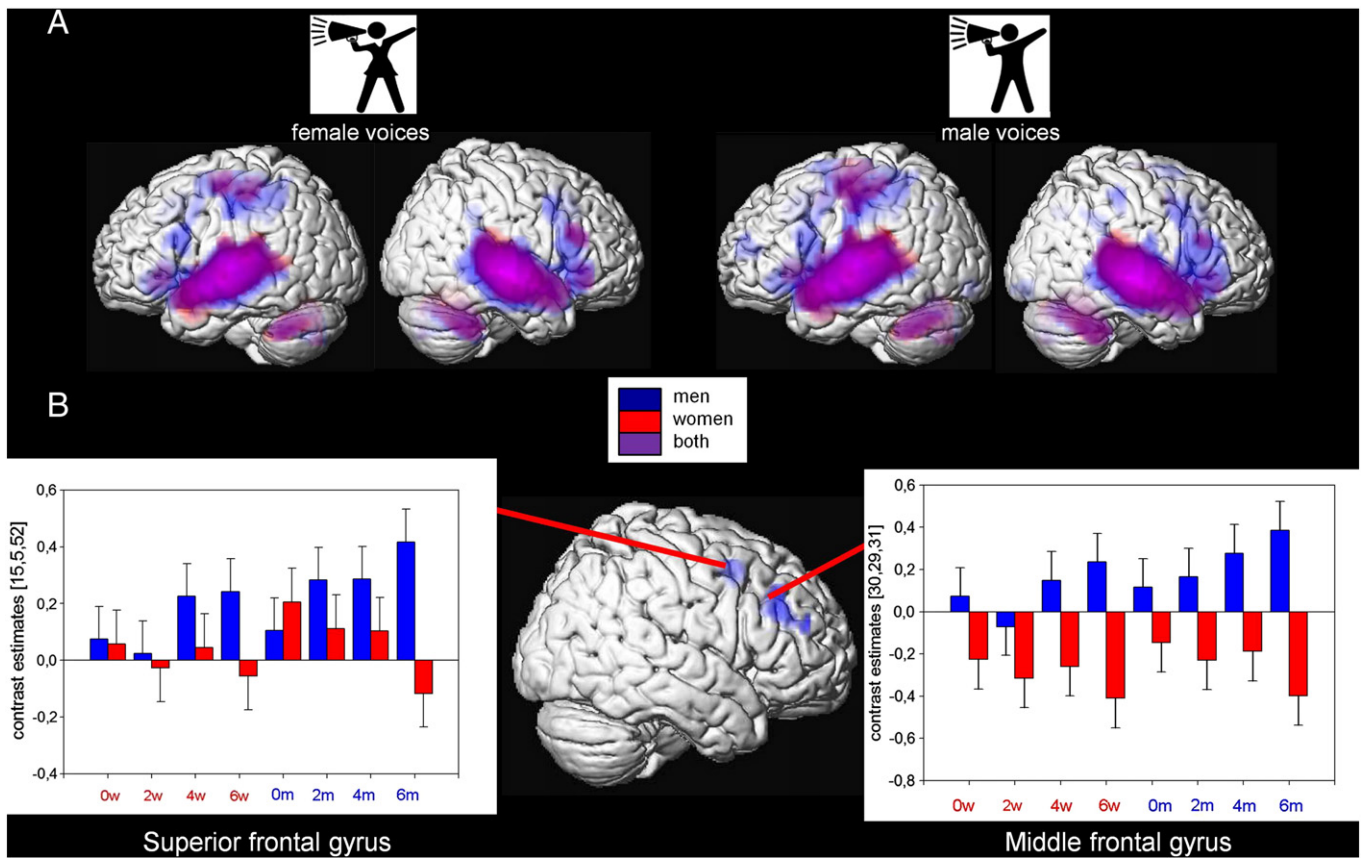


Fig. 5. A. Brain activation in men (blue) and women (red) separated for original male and female voices ($p < 0.05$ FWE corrected, extent threshold = 20 voxels) showing strong bilateral activation in typical voice related areas including superior temporal gyrus (note that purple colored regions represent activations shared by both sexes). B. Contrast estimates of stronger activation in the right superior and middle frontal gyrus in men (blue) as compared to women (red) for increasing morphing degree plotted for all 8 conditions ($p < 0.05$ Monte Carlo corrected).

(Drevets and Raichle, 1998). It integrates multiple pieces of information in order to choose a social behavior, in our case mate-identification, thereby accommodating mate values from multiple value dimensions (Funayama et al., 2012). Moreover, the MPFC is associated with monitoring of mental states of the self and seems to be necessary for self-referential thinking (Gusnard et al., 2001). Thus, brain activation in both prefrontal regions may also be of high relevance in the evaluation of potential mates.

Opposite-sex effect on temporo-parietal brain activation

Further, we found differential activation in response to female voices in the left posterior MTG and the angular gyrus in men compared to women as well as larger deactivation in response to female as compared to male voices in women. Both regions are meant to be multimodal, linked to auditory association cortex and involved in semantic processing (Binder et al., 2009). More detailed, the left

MTG and the angular gyrus are suggested to be involved in the linkage of acoustic input to the representation of a category with the best fit to a category being reflected by the least deactivation (Blumstein et al., 2005). Further, there is evidence from connectivity analyses that these two regions are highly connected to the MPFC when speech perception is requested and that this connection is strengthened when semantic context can positively influence speech perception (Obleser et al., 2007). Increased connectivity is also observed during the integration of facial and vocal information in person recognition (Joassin et al., 2011). The MTG and angular gyrus therefore seem to simultaneously direct attention to information from different sources to regulate social decision making.

In sum, from an evolutionary point of view, sex differences in voice gender perception may be related to higher attention to and emotional involvement as well as social reward of opposite sex information in the context of mate selection (Duncan et al., 2007), especially in men (Clutton-Brock and Vincent, 1991).

Sex differences in hierarchical voice processing

Voices are processed in a hierarchical manner from a perceptual and associative analysis in temporo-parietal cortices to its linkage to other modalities, memory content, social knowledge and attention in temporal as well as orbito- and middle prefrontal cortices (Blumstein et al., 2005; Christensen et al., 2008; Sabri et al., 2008; Warren et al., 2006).

We found no sex differences in activation in typical voice-selective auditory areas, such as the STG or the anterior temporal lobe. Instead we observed stronger activation/less deactivation in men compared

Table 4

Activation peaks (MNI coordinates) and cluster extension for the negative correlation of errors and brain activation in response to a linear increase of the voice morphing in women (with no positive correlation and no correlations in men; $p < .05$ FWE, cluster threshold > 20 voxels).

Brain region	Hemisphere	x	y	z	k	t
Insula	R	39	11	-14	58	6.27
Middle temporal gyrus	R	48	2	-29	31	5.64
Middle temporal pole	L	-33	14	-35	26	5.18
Paracentral lobe	L	-12	-19	76	22	5.06
Hippocampus	L	-15	-10	-20	20	4.54
Brain stem	R	3	-16	-20	50	4.51

to women for female vs. male voices in the left medial OFC, the MPFC and the left MTG extending to the angular gyrus. All three areas are associated with multimodal integration, decision making and directing attention especially in the social domain.

Correspondingly, a recent event-related potential (ERP) study showed a two stage voice gender processing with an early pitch processing potential around 50 ms associated with the auditory cortex and a relatively later gender differentiation at around 200 ms independent of pitch in fronto-central regions (Latinus and Taylor, 2012). Other studies also reported no activation increases in areas of early processing after the removal of variance related to acoustic parameters, such as pitch (Bestelmeyer et al., 2012; Wiethoff et al., 2008). Together with our findings during voice gender perception, evidence so far suggests sex differences occurring on higher order processing levels.

Increased activation in men compared to women in the superior temporal gyrus, i.e. at earlier processing stage, was found in a pitch perception task (Gaab et al., 2003) and may be related to the fact that the authors used sine wave tones (in contrast to voices), which lack social content. However, together with our findings this points to a stronger activity in men compared to women in several aspects of sound perception. Conflicting findings in this domain could be the consequence of methodological differences, such as a smaller statistical power due to small sample sizes possibly disguising subtle sex differences (e.g. 8 male and 8 female subjects in Lattner et al., 2005).

Interaction effects were mainly based on stronger down-regulation of the temporo-parietal junction, the MPFC and the OFC in response to female voices in women. In contrast, the only significant difference between female and male stimuli in men was found in the MPFC. It could be speculated that women tend to differentiate between male and female voices already in the association cortex and during the initial evaluation of emotional relevance (as reflected in stronger deactivation in the MTG and OFC in response to female voices) suggesting less relevance of and lower attention to same-sex voices. Moreover, stronger activation in response to female stimuli was found in men as compared to women. However, men seem to attach similar importance to same- and opposite sex voices on this rather early processing level. Yet, on a higher processing level, men also differentiate between the sexes as reflected by activation in the MPFC suggesting differences in the emotional attributions. In summary, increased MPFC activation in response to voices of the opposite sex was found in men and women.

However, this interpretation must be taken carefully, since post hoc results did not survive Bonferroni correction. Nevertheless and important to note, increased activation in response to female voices in men as compared to women as well as increased activation in earlier processing steps in response to male as compared to female voices in women seem to reflect found sex differences in behavioral performance, namely an opposite-sex effect in men and an opposite-sex bias in women. Further studies are needed to clarify the reliability of the results.

Gradual voice gender morphing

Voice gender morphing and cognitive processing load

We found notable performance differences between conditions indicating a general increase in cognitive demand with increasing stimulus modification in accordance with recent data (Charest et al., 2013). Task-associated d-primes revealed decreasing discrimination sensitivity with increasing morphing degree. However, male and female stimuli were distinguishable above chance level in all conditions except the highest morphing level (6st).

Brain activation changes in relation to an increasing gender morphing of the voices (comparable to Sokhi et al., 2005) demonstrated an effect of linear pitch and formant frequency scaling as illustrated by increasing activity in the right MCC. The cingulate cortex is associated with (phonological) conflict monitoring at a relatively high processing level (Haupt et al., 2009). The MCC is activated during

tasks on response selection and decision making (Vogt et al., 2003), the resolution of decision conflicts (Mitchell et al., 2009) and selective auditory attention (Sturm et al., 2011). It is linked to a mental inspection of alternative representations (Huijbers et al., 2011). In addition, there is evidence from brain lesion studies that damage in this area can lead to impaired voice discrimination (Hornak et al., 2003). Increased activation probably reflects additional cognitive demands due to higher task difficulty with increasing morphing level.

The notion is further affirmed by the fact that also activation in the IFG bilaterally and the left insula parametrically increases with increasing stimulus ambiguity (Fig. 2). Our data resemble recent results of Charest et al. (2013), who also found bilateral IFG and cingulate activation in response to increasing gender ambiguity in a voice perception task with a complex continuous carryover experimental design and a different voice morphing technique. Activation in the IFG and insula cortex were also associated with an increase in cognitive processing load due to semantic ambiguity (Rodd et al., 2010), increasing verbal working memory demands during the discrimination of linguistically relevant intonational features (LoCasto et al., 2004), sound categorization (Husain et al., 2006) and interference suppression (Vigneau et al., 2006). Activation in the bilateral IFG correlated positively with better performance in dichotic listening tasks (van Ettinger-Veenstra et al., 2010). However, using functional connectivity analysis Kemmotsu et al. (2005) showed that these activation increases may be less specific to language processing but rather related to more general selective attention processes. Enhanced activity in a network of IFG, MCC and insula was not only reported in response to unimodal auditory but also to tactile and visual detection tasks (Langner et al., 2012).

The anterior insula has already been shown to be involved in vocal sound processing (Wong et al., 2004), more exactly in the processing of vocal identity (Remedios et al., 2009) and the regulation of attentional listening demands (Christensen et al., 2008). Andics et al. (2010) reported reduced responses in the anterior insula when voice stimuli were most similar to a mean voice representation. This is in accordance with our finding of a relationship between insula activation increase with increasing morphing level and a decrease in errors during voice classification in women reflecting the central role of the anterior insula in the evaluation of vocal stimuli.

Voice gender morphing and sex differences

Comparing the sexes, men showed stronger brain activations than women in the superior and middle frontal gyrus. Especially the MFG is believed to be involved in focusing attention in speech perception and increases activation with cognitive load (Thomsen et al., 2004). Thus, men consulted more areas involved in higher cognitive processes than women to assess speakers' sex in voices parametrically becoming more ambiguous. An increase in effort in men was also reflected by reaction times, which (across the stimuli types of both genders) by trend were longer in men as compared to women. In addition, longer reaction times in combination with a lack of response bias may indicate that men continued to try to discriminate voices' gender. In contrast, women may shift to a response bias with increasing ambiguity resulting in decreased neural activation and shorter reaction times compared to men.

Finally, the negative correlation between error rates and activation in the left hippocampus in response to the morphing degree in women, but not in men, underscores the fact that the hippocampus is one of the evidently sexually dimorphic regions in the brain, differing in anatomical structure, neurochemistry and hormonal level (Cahill, 2006). The hippocampus is generally larger in women than in men when adjusted for total brain size (Goldstein et al., 2001). Functionally, it is a multimodal area involved in encoding and retrieval of associative information in memory especially for source information (Peters et al., 2007) and was found to have enhanced connectivity with auditory cortex (Joassin et al., 2011) suggesting

that its recruitment in voice discrimination in women may be a reflection of a comparison process of the stimuli with previous experiences, a strategy seemingly not comparably used in men. This indicates differential cognitive strategies in men and women.

Conclusion

Taken together, our findings support the notion of sex-specific brain networks enabling men and women to identify gender by means of vocal sounds. It seems that sex differences in the perception of male and female voices are reflected in networks of areas involved not only in auditory but also in attentional and evaluative processes.

Behaviorally, we could demonstrate an opposite-sex effect in voice perception in men and an opposite-sex bias in women. Brain activation in both sexes also points to an increased attention to voices of the opposite sex. Thereby, sex differences in neuronal response patterns can be localized on higher processing levels and reflect differences in behavioral outcome. All this indicates that these effects reflect sensitivity to the evolutionary relevance of voice perception. This is probably related to the central evolutionary role of voice interpretation in mate selection. The higher relevance of the opposite sex is reverberated by an additional recruitment of attention and interpretation-related areas.

By morphing original voices into an ambiguous gender range we could show additional sex differences namely increasing neuronal brain response in cognition related areas with increasing cognitive demands in men and a correlation between error rate and decrease of activation in medial temporal areas in women.

Neural sex differences during voice perception might act as confounding factors and have to be taken into account when investigating mixed sex samples in related imaging studies. Regarding future directions, the direct comparison of the explicit and implicit processing of gender-related information may be a promising research approach. Implicit information is processed comparatively subconsciously, which may be more comparable to normal social interaction.

Acknowledgments

The study was supported by the German Research Foundation (DFG: IRTG 1328 and HA 3202/7-1) and the Brain Imaging Facility of the Interdisciplinary Centre for Clinical Research within the faculty of Medicine at the RWTH Aachen University, Germany. The authors thank Carine Timma, Cordula Kemper and Thilo Kellermann for their assistance and support and all participants whose collaboration has made this study possible.

Conflict of interest

Authors declare no conflicts of interest.

References

- Aleman, A., Swart, M., 2008. Sex differences in neural activation to facial expressions denoting contempt and disgust. *PLoS One* 3, e3622.
- Andics, A., McQueen, J.M., Petersson, K.M., Gal, V., Rudas, G., Vidnyanszky, Z., 2010. Neural mechanisms for voice recognition. *NeuroImage* 52, 1528–1540.
- Apicella, C.L., Feinberg, D.R., 2009. Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proc. Biol. Sci.* 276, 1077–1082.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *NeuroImage* 26, 839–851.
- Baayen, R.H., Piepenbrock, R., Gulikers, L., 1995. The CELEX lexical database (CD-ROM). Linguistic Data Consortium University of Pennsylvania, Philadelphia, PA.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Belin, P., Zatorre, R.J., Ahad, P., 2002. Human temporal-lobe response to vocal sounds. *Brain Res. Cogn. Brain Res.* 13, 17–26.
- Belin, P., Fecteau, S., Bedard, C., 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn. Sci.* 8, 129–135.
- Belin, P., Bestelmeyer, P.E., Latinus, M., Watson, R., 2011. Understanding voice perception. *Br. J. Psychol.* 102, 711–725.
- Belizaire, G., Fillion-Bilodeau, S., Chartrand, J.P., Bertrand-Gauvin, C., Belin, P., 2007. Cerebral response to 'voiceness': a functional magnetic resonance imaging study. *Neuroreport* 18, 29–33.
- Bestelmeyer, P.E., Latinus, M., Bruckert, L., Rouger, J., Crabbe, F., Belin, P., 2012. Implicitly perceived vocal attractiveness modulates prefrontal cortex activity. *Cereb. Cortex* 22, 1263–1270.
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796.
- Blumstein, S.E., Myers, E.B., Rissman, J., 2005. The perception of voice onset time: an fMRI investigation of phonetic category structure. *J. Cogn. Neurosci.* 17, 1353–1366.
- Boersma, P., Weenink, D., 2010. Praat: doing phonetics by computer [Computer program]. Version 5.3.13, retrieved 11 April 2012 from <http://www.praat.org/>.
- Brown, E., Perrett, D.I., 1993. What gives a face its gender? *Perception* 22, 829–840.
- Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G.A., Kawahara, H., Belin, P., 2010. Vocal attractiveness increases by averaging. *Curr. Biol.* 20, 116–120.
- Brun, C.C., Lepore, N., Luders, E., Chou, Y.Y., Madsen, S.K., Toga, A.W., Thompson, P.M., 2009. Sex differences in brain structure in auditory and cingulate regions. *Neuroreport* 20, 930–935.
- Cahill, L., 2006. Why sex matters for neuroscience. *Nat. Rev. Neurosci.* 7, 477–484.
- Cellerino, A., Borghetti, D., Sartucci, F., 2004. Sex differences in face gender recognition in humans. *Brain Res. Bull.* 63, 443–449.
- Charest, I., Pernet, C., Latinus, M., Crabbe, F., Belin, P., 2013. Cerebral processing of voice gender studied using a continuous carryover fMRI design. *Cereb. Cortex* 23, 958–966.
- Christensen, T.A., Antonucci, S.M., Lockwood, J.L., Kittleson, M., Plante, E., 2008. Cortical and subcortical contributions to the attentive processing of speech. *Neuroreport* 19, 1101–1105.
- Clutton-Brock, T.H., Vincent, A.C., 1991. Sexual selection and the potential reproductive rates of males and females. *Nature* 351, 58–60.
- Conway, C.A., Jones, B.C., DeBruine, L.M., Little, A.C., 2008. Evidence for adaptive design in human gaze preference. *Proc. Biol. Sci.* 275, 63–69.
- Cornwell, R.E., Boothroyd, L., Burt, D.M., Feinberg, D.R., Jones, B.C., Little, A.C., Pitman, R., Whiten, S., Perrett, D.I., 2004. Concordant preferences for opposite-sex signals? Human pheromones and facial characteristics. *Proc. Biol. Sci.* 271, 635–640.
- Corso, J.F., 1963. Age and sex differences in pure-tone thresholds. Survey of hearing levels from 18 to 65 years. *Arch. Otolaryngol.* 77, 385–405.
- Cox, R.W., 2012. AFNI: what a long strange trip it's been. *Neuroimage* 62, 743–747.
- Dacakis, G., 2000. Long-term maintenance of fundamental frequency increases in male-to-female transsexuals. *J. Voice* 14, 549–556.
- D'Ausilio, A., Bufalari, I., Salmas, P., Busan, P., Fadiga, L., 2011. Vocal pitch discrimination in the motor system. *Brain Lang.* 118, 9–14.
- Davey, C.G., Allen, N.B., Harrison, B.J., Dwyer, D.B., Yucel, M., 2010. Being liked activates primary reward and midline self-related brain regions. *Hum. Brain Mapp.* 31, 660–668.
- Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D.I., Schneider, F., Habel, U., 2010. Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology* 35, 67–82.
- Drevets, W.C., Raichle, M.E., 1998. Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive process: implications for interaction between cognition and emotion. *Cogn. Emot.* 12, 353–385.
- Duncan, L.A., Park, J.H., Faulkner, J., Schaller, M., Neuberg, S.L., Kenrick, D.T., 2007. Adaptive allocation of attention: effects of sex and sociosexuality on visual attention to attractive opposite-sex faces. *Evol. Hum. Behav.* 28, 359–364.
- Fant, G., 1960. *Acoustic Theory of Speech Production*. Mouton & Co. N.V., Publishers, The Hague.
- Feinberg, D.R., 2008. Are human faces and voices ornaments signaling common underlying cues to mate value? *Evol. Anthropol.* 17, 112–118.
- 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.
- Fischer, H., Sandblom, J., Herlitz, A., Fransson, P., Wright, C.I., Backman, L., 2004. Sex-differential brain activation during exposure to female and male faces. *Neuroreport* 15, 235–238.
- Fitch, W.T., 2000. The evolution of speech: a comparative review. *Trends Cogn. Sci.* 4, 258–267.
- Fitch, W.T., Giedd, J., 1999. Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoust. Soc. Am.* 106, 1511–1522.
- Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science* 322, 970–973.
- Friedman, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* 32, 675–701.
- Funayama, R., Sugiura, M., Sassa, Y., Jeong, H., Wakusawa, K., Horie, K., Sato, S., Kawashima, R., 2012. Neural bases of human mate choice: multiple value dimensions, sex difference, and self-assessment system. *Soc. Neurosci.* 7, 59–73.
- Gaab, N., Keenan, J.P., Schlaug, G., 2003. The effects of gender on the neural substrates of pitch memory. *J. Cogn. Neurosci.* 15, 810–820.
- Gelfer, M.P., Schofield, K.J., 2000. Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *J. Voice* 14, 22–33.
- Ghazanfar, A.A., Rendall, D., 2008. Evolution of human vocal production. *Curr. Biol.* 18, R457–R460.
- Goldstein, J.M., Seidman, L.J., Horton, N.J., Makris, N., Kennedy, D.N., Caviness Jr., V.S., Faraone, S.V., Tsuang, M.T., 2001. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb. Cortex* 11, 490–497.

- Goldstein, J.M., Jerram, M., Abbs, B., Whitfield-Gabrieli, S., Makris, N., 2010. Sex differences in stress response circuitry activation dependent on female hormonal cycle. *J. Neurosci.* 30, 431–438.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4259–4264.
- Haupt, S., Axmacher, N., Cohen, M.X., Elger, C.E., Fell, J., 2009. Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Hum. Brain Mapp.* 30, 3043–3056.
- Hauser, M., 1996. *The evolution of communication*. MIT Press, Cambridge, MA.
- Hillenbrand, J.M., Clark, M.J., 2009. The role of $f(0)$ and formant frequencies in distinguishing the voices of men and women. *Atten. Percept. Psychophys.* 71, 1150–1166.
- Hodges-Simeon, C.R., Gaulin, S.J., Puts, D.A., 2010. Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* 21, 406–427.
- Hodges-Simeon, C.R., Gaulin, S.J., Puts, D.A., 2011. Voice correlates of mating success in men: examining “contests” versus “mate choice” modes of sexual selection. *Arch. Sex. Behav.* 40, 551–557.
- Hofmann, S.G., Suvak, M., Litz, B.T., 2006. Sex differences in face recognition and influence of facial affect. *Personal. Individ. Differ.* 40, 1683–1690.
- Honorof, D.N., Whalen, D.H., 2010. Identification of speaker sex from one vowel across a range of fundamental frequencies. *J. Acoust. Soc. Am.* 128, 3095–3104.
- Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O’Doherty, J., Bullock, P.R., Polkey, C.E., 2003. Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain* 126, 1691–1712.
- Hughes, S.M., Pastizzo, M.J., Gallup, G.G., 2008. The sound of symmetry revisited: subjective and objective analyses of voices. *J. Nonverbal Behav.* 32, 93–108.
- Huijbers, W., Pennartz, C.M., Rubin, D.C., Daselaar, S.M., 2011. Imagery and retrieval of auditory and visual information: neural correlates of successful and unsuccessful performance. *Neuropsychologia* 49, 1730–1740.
- Husain, F.T., McKinney, C.M., Horwitz, B., 2006. Frontal cortex functional connectivity changes during sound categorization. *Neuroreport* 17, 617–621.
- Ishai, A., 2007. Sex, beauty and the orbitofrontal cortex. *Int. J. Psychophysiol.* 63, 181–185.
- Joassin, F., Pesenti, M., Maurage, P., Verreault, E., Bruyer, R., Campanella, S., 2011. Cross-modal interactions between human faces and voices involved in person recognition. *Cortex* 47, 367–376.
- 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. *Biol. Lett.* 4, 192–194.
- Jones, B., 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. *Anim. Behav.* 79, 57–62.
- Kemmotsu, N., Villalobos, M.E., Gaffrey, M.S., Courchesne, E., Muller, R.A., 2005. Activity and functional connectivity of inferior frontal cortex associated with response conflict. *Brain Res. Cogn. Brain Res.* 24, 335–342.
- Koch, K., Pauly, K., Kellermann, T., Seiferth, N.Y., Reske, M., Backes, V., Stocker, T., Shah, N.J., Amunts, K., Kircher, T., Schneider, F., Habel, U., 2007. Gender differences in the cognitive control of emotion: an fMRI study. *Neuropsychologia* 45, 2744–2754.
- Kohn, N., Kellermann, T., Gur, R.C., Schneider, F., Habel, U., 2011. Gender differences in the neural correlates of humor processing: implications for different processing modes. *Neuropsychologia* 49, 888–897.
- Krach, S., Blumel, I., Marjoram, D., Lataster, T., Krabbendam, L., Weber, J., van Os, J., Kircher, T., 2009. Are women better mindreaders? Sex differences in neural correlates of mentalizing detected with functional MRI. *BMC Neurosci.* 10, 9.
- Krizman, J., Skoe, E., Kraus, N., 2012. Sex differences in auditory subcortical function. *Clin. Neurophysiol.* 123, 590–597.
- Langner, R., Kellermann, T., Eickhoff, S.B., Boers, F., Chatterjee, A., Willmes, K., Sturm, W., 2012. Staying responsive to the world: modality-specific and -nonspecific contributions to speeded auditory, tactile, and visual stimulus detection. *Hum. Brain Mapp.* 33, 398–418.
- Latinus, M., Belin, P., 2012. Perceptual auditory aftereffects on voice identity using brief vowel stimuli. *PLoS One* 7, e41384.
- Latinus, M., Taylor, M.J., 2012. Discriminating male and female voices: differentiating pitch and gender. *Brain Topogr.* 25, 194–204.
- Latinus, M., Crabbe, F., Belin, P., 2011. Learning-induced changes in the cerebral processing of voice identity. *Cereb. Cortex* 21, 2820–2828.
- Lattner, S., Friederici, A.D., 2003. Talker’s voice and gender stereotype in human auditory sentence processing—evidence from event-related brain potentials. *Neurosci. Lett.* 339, 191–194.
- Lattner, S., Meyer, M.E., Friederici, A.D., 2005. Voice perception: sex, pitch, and the right hemisphere. *Hum. Brain Mapp.* 24, 11–20.
- Lehmann, M., Lehnert, B., Neuschaefer-Rube, C., 2010. Vergleichende Untersuchungen zur Geschlechtererkennung von Sprechern auf Einzellaute-, Silben-, Wort- und Satzebene. *Jahresversammlung der Deutschen Gesellschaft für Hals-Nasen-Ohren-Heilkunde, Kopf- und Hals-Chirurgie e.V.*, Wiesbaden, 81.
- Little, A.C., Jones, B.C., 2006. Attraction independent of detection suggests special mechanisms for symmetry preferences in human face perception. *Proc. Biol. Sci.* 273, 3093–3099.
- LoCasto, P.C., Krebs-Noble, D., Gullapalli, R.P., Burton, M.W., 2004. An fMRI investigation of speech and tone segmentation. *J. Cogn. Neurosci.* 16, 1612–1624.
- Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18, 50–60.
- Mather, M., Lighthall, N.R., Nga, L., Gorlick, M.A., 2010. Sex differences in how stress affects brain activity during face viewing. *Neuroreport* 21, 933–937.
- McNally, R.J., Otto, M.W., Hornig, C.D., 2001. The voice of emotional memory: content-filtered speech in panic disorder, social phobia, and major depressive disorder. *Behav. Res. Ther.* 39, 1329–1337.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Mitchell, D.G., Luo, Q., Avny, S.B., Kasprzycki, T., Gupta, K., Chen, G., Finger, E.C., Blair, R.J., 2009. Adapting to dynamic stimulus–response values: differential contributions of inferior frontal, dorsomedial, and dorsolateral regions of prefrontal cortex to decision making. *J. Neurosci.* 29, 10827–10834.
- Mullenix, J.W., Johnson, K.A., Topcu-Durgun, M., Farnsworth, L.M., 1995. The perceptual representation of voice gender. *J. Acoust. Soc. Am.* 98 (6), 3080–3095.
- Narins, P.M., Capranica, R.R., 1976. Sexual differences in the auditory system of the tree frog *Eleutherodactylus coqui*. *Science* 192, 378–380.
- Oates, J., Dacakis, G., 1997. Voice change in transsexuals. *Venereology* 10, 178–187.
- Obleser, J., Wise, R.J., Alex Dresner, M., Scott, S.K., 2007. Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* 27, 2283–2289.
- O’Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J., 2003. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41, 147–155.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- O’Toole, A.J., Deffenbacher, K.A., Valentin, D., McKee, K., Huff, D., Abdi, H., 1998. The perception of face gender: the role of stimulus structure in recognition and classification. *Mem. Cognit.* 26, 146–160.
- Pernet, C.R., Belin, P., 2012. The role of pitch and timbre in voice gender categorization. *Front. Psychol.* 3, 23.
- Peters, J., Suchan, B., Koster, O., Daum, I., 2007. Domain-specific retrieval of source information in the medial temporal lobe. *Eur. J. Neurosci.* 26, 1333–1343.
- Phillips, M.L., Ladouceur, C.D., Drevets, W.C., 2008. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol. Psychiatry* 13 (829), 833–857.
- Poon, S., Ng, M.L., 2011. Contributions of voice fundamental frequency and formants to the identification of speaker’s gender. *Proceedings of the 17th International Congress of Phonetic Sciences*.
- Proverbio, A.M., Riva, F., Martin, E., Zani, A., 2010. Neural markers of opposite-sex bias in face processing. *Front. Psychol.* 1, 169.
- Rademacher, J., Morosan, P., Schleicher, A., Freund, H.J., Zilles, K., 2001. Human primary auditory cortex in women and men. *Neuroreport* 12, 1561–1565.
- Ray, R.D., Zald, D.H., 2012. Anatomical insights into the interaction of emotion and cognition in the prefrontal cortex. *Neurosci. Biobehav. Rev.* 36, 479–501.
- Remedios, R., Logothetis, N.K., Kayser, C., 2009. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J. Neurosci.* 29, 1034–1045.
- Rendall, D., Owren, M.J., Weerts, E., Hienz, R.D., 2004. Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. *J. Acoust. Soc. Am.* 115, 411–421.
- Rodd, J.M., Johnsrude, I.S., Davis, M.H., 2010. The role of domain-general frontal systems in language comprehension: evidence from dual-task interference and semantic ambiguity. *Brain Lang.* 115, 182–188.
- Rodero, E., 2011. Intonation and emotion: influence of pitch levels and contour type on creating emotions. *J. Voice* 25, e25–e34.
- Ruytjens, L., Georgiadis, J.R., Holstege, G., Wit, H.P., Albers, F.W., Willemsen, A.T., 2007. Functional sex differences in human primary auditory cortex. *Eur. J. Nucl. Med. Mol. Imaging* 34, 2073–2081.
- Ryan, M.J., Perrill, S.A., Wilczynski, W., 1992. Auditory tuning and call frequency predict population-based mating preferences in the cricket frog, *Acris crepitans*. *Am. Nat.* 139, 1370–1383.
- Sabatinelli, D., Flaisch, T., Bradley, M.M., Fitzsimmons, J.R., Lang, P.J., 2004. Affective picture perception: gender differences in visual cortex? *Neuroreport* 15, 1109–1112.
- Sabri, M., Binder, J.R., Desai, R., Medler, D.A., Leit, M.D., Lieberthal, E., 2008. Attentional and linguistic interactions in speech perception. *Neuroimage* 39, 1444–1456.
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., Vuilleumier, P., 2005. Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *Neuroimage* 28, 848–858.
- Sato, H., Sando, I., Takahashi, H., 1991. Sexual dimorphism and development of the human cochlea. *Computer 3-D measurement. Acta Otolaryngol.* 111, 1037–1040.
- Schweinberger, S.R., Casper, C., Hauthal, N., Kaufmann, J.M., Kawahara, H., Kloth, N., Robertson, D.M., Simpson, A.P., Zanke, R., 2008. Auditory adaptation in voice perception. *Curr. Biol.* 18, 684–688.
- Schweinberger, S.R., Walther, C., Zanke, R., Kovacs, G., 2011. Neural correlates of adaptation to voice identity. *Br. J. Psychol.* 102, 748–764.
- Shirao, N., Okamoto, Y., Mantani, T., Yamawaki, S., 2005. Gender differences in brain activity generated by unpleasant word stimuli concerning body image: an fMRI study. *Br. J. Psychiatry* 186, 48–53.
- Simpson, A.P., 2009. Phonetic differences between male and female speech. *Lang. Linguist. Compass* 3, 621–640.
- Sokhi, D.S., Hunter, M.D., Wilkinson, I.D., Woodruff, P.W., 2005. Male and female voices activate distinct regions in the male brain. *Neuroimage* 27, 572–578.
- Spencer, L.E., 1988. Speech characteristics of male-to-female transsexuals: a perceptual and acoustic study. *Folia Phoniatr. (Basel)* 40, 31–42.
- Speckelmeier, K.N., Rademacher, L., Paulus, F.M., Grunder, G., 2012. Neural activation during anticipation of opposite-sex and same-sex faces in heterosexual men and women. *Neuroimage* 66C, 223–231.
- Sturm, W., Schnitker, R., Grande, M., Huber, W., Willmes, K., 2011. Common networks for selective auditory attention for sounds and words? An fMRI study with implications for attention rehabilitation. *Restor. Neurol. Neurosci.* 29, 73–83.
- Thomsen, T., Rimol, L.M., Ersland, L., Hugdahl, K., 2004. Dichotic listening reveals functional specificity in prefrontal cortex: an fMRI study. *Neuroimage* 21, 211–218.

- Titze, I.R., 1989. Physiologic and acoustic differences between male and female voices. *J. Acoust. Soc. Am.* 85, 1699–1707.
- Titze, I.R., 2008. Nonlinear source-filter coupling in phonation: theory. *J. Acoust. Soc. Am.* 123, 2733–2749.
- Tovee, M.J., Maisey, D.S., Emery, J.L., Cornelissen, P.L., 1999. Visual cues to female physical attractiveness. *Proc. Biol. Sci.* 266, 211–218.
- van Dommelen, W.A., 1990. Acoustic parameters in human speaker recognition. *Lang. Speech* 33 (Pt 3), 259–272.
- van Ettinger-Veenstra, H.M., Ragnehed, M., Hallgren, M., Karlsson, T., Landtblom, A.M., Lundberg, P., Engstrom, M., 2010. Right-hemispheric brain activation correlates to language performance. *Neuroimage* 49, 3481–3488.
- Vigneau, M., Beaucousin, V., Herve, P.Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432.
- Vogt, B.A., Berger, G.R., Derbyshire, S.W., 2003. Structural and functional dichotomy of human midcingulate cortex. *Eur. J. Neurosci.* 18, 3134–3144.
- von Kriegstein, K., Eger, E., Kleinschmidt, A., Giraud, A.L., 2003. Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Res. Cogn. Brain Res.* 17, 48–55.
- von Kriegstein, K., Smith, D.R., Patterson, R.D., Kiebel, S.J., Griffiths, T.D., 2010. How the human brain recognizes speech in the context of changing speakers. *J. Neurosci.* 30, 629–638.
- Wallentin, M., 2009. Putative sex differences in verbal abilities and language cortex: a critical review. *Brain Lang.* 108, 175–183.
- Ward, B.D., 2000. Simultaneous inference for fMRI data. AFNI 3dDeconvolve Documentation. Medical College of Wisconsin, Milwaukee, WI.
- Warren, J.D., Scott, S.K., Price, C.J., Griffiths, T.D., 2006. Human brain mechanisms for the early analysis of voices. *Neuroimage* 31, 1389–1397.
- Wickens, T., 2002. *Elementary Signal Detection Theory*. Oxford University Press, Los Angeles.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., Ethofer, T., 2008. Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. *Neuroimage* 39, 885–893.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biom. Bull.* 1, 80–83.
- Wittchen, H.U., Zaudig, M., Fydrich, T., 1997. *Strukturiertes Klinisches Interview für DSM-IV*. Hogrefe, Göttingen.
- Wong, P.C., Parsons, L.M., Martinez, M., Diehl, R.L., 2004. The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. *J. Neurosci.* 24, 9153–9160.
- Yamaguchi, M.K., Hirukawa, T., Kanazawa, S., 1995. Judgment of gender through facial parts. *Perception* 24, 563–575.
- Zaske, R., Schweinberger, S.R., Kaufmann, J.M., Kawahara, H., 2009. In the ear of the beholder: neural correlates of adaptation to voice gender. *Eur. J. Neurosci.* 30, 527–534.