



Ahrens, Merle M., Hasan, Bashar Awwad Shiekh, Giordano, Bruno, and Belin, Pascal (2014) Gender differences in the temporal voice areas. *Frontiers in Neuroscience* . ISSN 1662-4548

Copyright © 2014 The Authors

<http://eprints.gla.ac.uk/95016>

Deposited on: 11 July 2014

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Gender differences in the temporal voice areas

Merle Marie Ahrens, Bashar Awwad_Shiekh_Hasan, Bruno L. Giordano and Pascal Belin

Journal Name:	Frontiers in Neuroscience
ISSN:	1662-453X
Article type:	Original Research Article
Received on:	09 Dec 2013
Accepted on:	10 Jul 2014
Provisional PDF published on:	10 Jul 2014
www.frontiersin.org:	www.frontiersin.org
Citation:	Ahrens MM, Awwad_shiekh_hasan B, Giordano BL and Belin P(2014) Gender differences in the temporal voice areas. <i>Front. Neurosci.</i> 8:228. doi:10.3389/fnins.2014.00228
Copyright statement:	© 2014 Ahrens, Awwad_shiekh_hasan, Giordano and Belin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY) . The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

This Provisional PDF corresponds to the article as it appeared upon acceptance, after rigorous peer-review. Fully formatted PDF and full text (HTML) versions will be made available soon.

Gender differences in the temporal voice areas

Merle-Marie Ahrens^{1,2*}, Bashar Awwad Shiekh Hasan¹, Bruno L. Giordano¹, Pascal Belin^{1,2,3}

¹ Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK

² Institut des Neurosciences de La Timone, UMR 7289, CNRS & Université Aix-Marseille, Marseille, France

³ International Laboratories for Brain, Music and Sound, Université de Montréal and McGill University, Montreal, QC, Canada

Correspondence

Merle-Marie Ahrens
University of Glasgow
Institute of Neuroscience and Psychology
58 Hillhead Street
Glasgow G12 8QB, UK
merlea@psy.gla.ac.uk

Abstract

There is not only evidence for behavioral differences in voice perception between female and male listeners, but also recent suggestions for differences in neural correlates between genders. The fMRI functional voice localizer (comprising a univariate analysis contrasting stimulation with vocal versus non-vocal sounds) is known to give robust estimates of the temporal voice areas (TVAs). However, there is growing interest in employing multivariate analysis approaches to fMRI data (e.g. multivariate pattern analysis; MVPA). The aim of the current study was to localize voice-related areas in both female and male listeners and to investigate whether brain maps may differ depending on the gender of the listener. After a univariate analysis, a random effects analysis was performed on female ($n = 149$) and male ($n = 123$) listeners and contrasts between them were computed. In addition, MVPA with a whole-brain searchlight approach was implemented and classification maps were entered into a second-level permutation based random effects models using statistical non-parametric mapping (SnPM; Nichols & Holmes 2002). Gender differences were found only in the MVPA. Identified regions were located in the middle part of the middle temporal gyrus (bilateral) and the middle superior temporal gyrus (right hemisphere). Our results suggest differences in classifier performance between genders in response to the voice localizer with higher classification accuracy from local BOLD signal patterns in several temporal-lobe regions in female listeners.

Keywords: Gender difference, fMRI, voice localizer, temporal voice areas, multivariate pattern analysis (MVPA), voice perception

40 Introduction

41 Prior functional magnetic resonance imaging (fMRI) findings suggest a robust brain response
42 to vocal vs. non-vocal sounds **in many regions of the human auditory cortex in particular in**
43 **the superior temporal gyrus (STG).** Vocal sounds, including but not restricted to speech
44 sounds, evoke a greater response than non-vocal sounds with bilateral activation foci located
45 near the anterior part of the STG extending to anterior parts of the superior temporal sulcus
46 and posterior foci located in the middle superior temporal sulcus (STS) (Belin et al., 2002,
47 2000; Binder et al., 2000). Using the functional voice localizer, these findings were replicated
48 and used in various studies (Belin et al., 2002, 2000; Ethofer et al., 2012; Kreifelts et al.,
49 2009; Latinus et al., 2011). The conventional way of identifying voice sensitive regions is by
50 applying univariate statistics, implemented using a Generalized-Linear Model (GLM), to
51 fMRI data assuming independence among voxels.

52 Interest has recently grown in applying multivariate approaches (e.g. Multivariate pattern
53 analysis; MVPA). Instead of modeling individual voxels independently (univariate analysis),
54 MVPA considers the information of distributed pattern in several voxels (e.g. Mur et al.,
55 2009; Norman et al., 2006). Several studies used multivariate approaches to decode
56 information reflected in brain activity patterns related to specific experimental conditions
57 (Cox & Savoy, 2003; Haynes & Rees, 2005, 2006; Kotz et al., 2013). MVPA is usually
58 applied on unsmoothed data preserving high spatial frequency information. Thus, MVPA is
59 argued to be more sensitive in detecting different cognitive states. In contrast, the
60 conventional univariate analysis averages across voxels, thereby removing focally distributed
61 effects (spatial smoothing). The smoothing across voxels may lead to a reduction in the
62 information content (Haynes et al., 2007; Kriegeskorte et al., 2006; Norman et al., 2006). At
63 present, a multivariate approach has never been employed to investigate whether it may yield
64 a different pattern of voice-specific (voice/non-voice classification) brain regions compared
65 to the univariate analysis.

66 The voice contains socially and biologically relevant information and plays a crucial role in
67 human interaction. This information is particularly relevant for interaction between different
68 genders (e.g. regarding emotions, identities, attractiveness) (Belin et al., 2011, 2004). Overall,
69 research suggests that women are more sensitive than men in emotion recognition from faces
70 and voices (J. A. Hall et al., 2006; J. A. Hall, 1978; Schirmer & Kotz, 2006). Women perform
71 better in judging others' nonverbal behavior (J. A. Hall, 1978) and seem to process nonverbal
72 emotional information more automatically as compared to men (Schirmer et al., 2005). In
73 addition, women but not men show greater limbic activity when processing emotional facial
74 expressions (G. Hall et al., 2004). The exact neural mechanisms underlying voice processing
75 in both female and male listeners still remains under debate. For instance, a study by Lattner
76 et al. (2005) found no significant difference between the activation patterns of female and
77 male listeners in response to voice-related information. However, there is evidence from both
78 behavioral and neural activation studies for differences in voice perception between listeners'
79 gender (Junger et al., 2013; Schirmer et al., 2002, 2007, 2004; Shaywitz et al., 1995; Skuk &
80 Schweinberger, 2013).

81 A recent behavioral study by Skuk and Schweinberger (2013) investigated gender differences
82 in a familiar voice identification task. They found an own-gender bias for males but not for
83 females while females outperformed males overall. These behavioral differences **(Skuk &**
84 **Schweinberger, 2013)** may also be reflected by differences in neural activity. Previous fMRI
85 studies investigating potential neural correlates suggested a sex difference in the functional
86 organization of the brain for phonological processing (Shaywitz et al., 1995), in emotional

87 prosodic and semantic processing (Schirmer et al., 2002, 2004) and in response to gender-
88 specific voice perception (Junger et al., 2013). Further evidence suggests differences between
89 genders in vocal processing shown by an EEG study, where the processing of vocal sounds
90 with more emotional and/or social information was more sensitive in women as compared to
91 men (Schirmer & Kotz, 2006; Schirmer et al., 2007). The above-mentioned studies mainly
92 focus on gender differences in emotional speech processing or opposite-sex perception.
93 However, identified brain regions are not consistent: different experimental designs and
94 applied methods vary and make it difficult to compare between these studies (Junger et al.,
95 2013; Schirmer et al., 2002, 2007, 2004; Shaywitz et al., 1995).

96 The current study employs a well-established experimental design of the functional ‘voice
97 localizer’, known to give robust estimates of the TVAs across the majority of participants.
98 The voice localizer includes a variety of different vocal sounds, not exclusively female or
99 male voices, but also speech and non-speech of women, men and infants and non-vocal
100 sounds (e.g. environmental sounds). In this study, we were interested in the effect of gender
101 on the results of the voice localizer and we asked an explorative research question of whether
102 brain activation and/or classification accuracy maps in response to vocal (speech and non-
103 speech) and non-vocal sounds differ between female and male listeners without prior
104 assumptions about the strength of voice-specific activity.

105 The voice localizer paradigm is often used in the literature (Belin et al., 2002, 2000; Ethofer
106 et al., 2012; Kreifelts et al., 2009; Latinus et al., 2011), which makes it easier to compare
107 among studies as well as among participants or groups. Instead of using the conventional
108 univariate method, employing MVPA may offer a more sensitive approach in order to study
109 potential differences between genders by means of above chance vocal/non-vocal
110 classification accuracies in different regions of the brain. Therefore, we investigated our
111 research question by implementing the conventional univariate analysis using GLM and
112 MVPA based on a support-vector machine (SVM) classifier with a spherical searchlight
113 approach. This approach enabled us to explore cortical activity over the whole-brain and to
114 examine whether activation and/or classification maps in response to the voice localizer may
115 significantly differ between genders. Since the effect size between genders is expected to be
116 very small, the current study offers a substantially large sample size with $n = 149$ females and
117 $n = 123$ males. Thus, this study provides a large sample size, a well-established experimental
118 design and the direct comparison of two different fMRI data analysis approaches applied on
119 the exact same data.

120 **Methods**

121 **Participants**

122 fMRI data of 272 healthy participants, 149 female (age range: 18-68 years; mean \pm s.d. =
123 24.5 ± 8.0) and 123 male (age range: 18-61 years; mean \pm s.d. = 24.4 ± 6.5) with self-
124 reported normal audition were analyzed. This study was conducted at the Institute of
125 Neuroscience and Psychology (INP) in Glasgow and approved by the ethics committee of the
126 University of Glasgow. Volunteers provided written informed consent before participating
127 and were paid afterwards.

128 **Voice localizer paradigm**

129 Subjects were instructed to close their eyes and passively listen to a large variety of sounds.
130 Stimuli were presented in a simple block design and divided into vocal (20 blocks) and non-
131 vocal (20 blocks) conditions. **Vocal blocks contained only sounds of human vocal origin**

132 (excluding sounds without vocal fold vibration such as whistling or whispering) and
133 consisted of speech (e.g. words, syllables, connected speech in different languages) or non-
134 speech (e.g. coughs, laughs, sighs and cries). The vocal stimuli consisted of recordings from
135 7 babies, 12 adults, 23 children, and 5 elderly people. Half of the vocal sounds (speech and
136 non-speech) consisted of vocalizations from adults and elderly people (women and men) with
137 comparable proportions for both genders (~24% female, ~22% male). The other half of the
138 vocal sounds consisted of infant vocalizations (speech and non-speech) which also included
139 baby crying/laughing. Recorded non-vocal sounds included various environmental sounds
140 (e.g. animal vocalizations, musical instruments, nature and industrial sounds). A total number
141 of 40 blocks were presented. Each block lasted for 8 seconds with an inter-block interval of
142 2 seconds. Stimuli (16bit, mono, 22050 Hz sampling rate) were normalized for RMS and are
143 available at <http://vnl.psy.gla.ac.uk/resources.php> (Belin et al., 2000).

144 **MRI data acquisition**

145 Scanning was carried out in a 3T MR scanner (Magnetom Trio Siemens, Erlangen, Germany)
146 and all data were acquired with the same scanner at the INP in Glasgow. Functional MRI
147 volumes of the whole cortex were acquired using an **echo-planar gradient pulse sequence**
148 (voxel size = 3 mm x 3 mm x 3 mm; Time of Repetition (TR) = 2000 ms; Echo Time (TE) =
149 30 ms; slice thickness = 3 mm; inter-slice gap = 0.3 mm; field of view (FoV) = 210 mm;
150 matrix size = 70 x 70; excitation angle = 77°). A total number of 310 volumes (32 slices per
151 volume, interleaved acquisition order) were collected with a total acquisition time of
152 10.28 minutes. Anatomical MRI volumes were acquired using a **magnetization-prepared**
153 **rapid gradient echo sequence (MPRAGE)** (voxel size = 1 mm x 1 mm x 1 mm; TR =
154 1900 ms; TE = 2.52 ms; inversion time (TI) = 900 ms; slice thickness = 1 mm; FoV =
155 256 mm; matrix size = 256 x 265; excitation angle = 9°; 192 axial slices).

156 **fMRI data analysis**

157 **Pre-processing**

158 Pre-processing was performed using the statistical parametric mapping software SPM8
159 (Department of Cognitive Neurology, London, UK.
160 <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). After reorientation of functional and
161 anatomical volumes to the AC/PC line (anterior- and posterior commissure), functional
162 images were motion corrected (standard realignment). Since, subjects may have moved
163 between anatomical and functional data acquisition, the anatomical volumes were co-
164 registered to the mean functional image produced in the realignment above. Anatomical
165 volumes were segmented in order to generate a binary gray matter template at threshold
166 probability level of 0.5 for each individual participant. This template was applied during
167 model specification in both univariate analysis und MVPA. For the univariate processing,
168 realigned functional volumes were normalized to a standard MNI template (Montreal
169 Neurological Institute) and spatially smoothed with a 6 mm full-width at half mean (FWHM)
170 Gaussian Kernel.

171 **Univariate analysis**

172 The design matrix was defined such that each block of the experimental paradigm correlated
173 to one condition, yielding a design matrix with 20 onsets for each condition (vocal and non-
174 vocal). Analysis was based on the conventional general linear model (GLM) and stimuli were
175 convolved with a boxcar hemodynamic response function provided by SPM8. Contrast
176 images of vocal versus non-vocal conditions were generated for each individual subject and
177 entered into a second-level random effects analysis (RFX). To declare at the group-level
178 whether any difference between the two conditions was significantly larger than zero, a one-

179 sample t-test was applied and FWE-corrected ($p < 0.05$) brain maps were calculated. To
180 investigate whether brain activity significantly differs between genders in response to vocal
181 versus non-vocal sounds, contrasts between females versus males (male > female, female >
182 male) were computed in a second level RFX analysis (two-sample t-test; $p < 0.05$ FWE-
183 corrected). This analysis was restricted to voxels with classification accuracy significantly
184 above theoretical chance ($p < 0.01$ uncorrected) in both females and males (see multivariate
185 pattern analysis below and yellow area in Fig. 2).

186 **Multivariate pattern analysis**

187 Multivariate pattern classification was performed on unsmoothed and non-normalized data
188 using Matlab (Mathworks Inc., Natick, USA) and in-house utility scripts (INP, Voice
189 Neurocognition Laboratory; Dr. Bashar Awwad Shiekh Hasan and Dr. Bruno L. Giordano),
190 where the default linear support vector machine (SVM) classifier was applied. The classifier
191 was trained and separately tested following a leave-one out cross validation strategy applied
192 on the 40 beta parameter estimates obtained from the univariate analysis (GLM).

193 A whole-brain searchlight decoding analysis was implemented using a sphere with a radius of
194 6 mm (average number of voxels in one sphere: 20.6 ± 1.0 s.d.) (Kriegeskorte et al., 2006). A
195 sphere was only considered for analysis if a minimum of 50% of its voxels were within the
196 gray matter. The data of the voxels within a sphere were classified and the classification
197 accuracy was stored at the central voxel, yielding a 3D brain map of classification accuracy
198 (percentage of correct classifications) (Haynes et al., 2007; Kriegeskorte et al., 2006). To
199 identify brain regions in which classification accuracy was significantly above chance by
200 females and males, the theoretical chance level (50%) was subtracted then normalized (to the
201 MNI template) and smoothed (6 mm FWHM Gaussian Kernel). To make inference on female
202 and male participants, classification brain maps were entered into a second-level permutation
203 based analysis using statistical nonparametric mapping (SnPM; Statistical NonParametric
204 Mapping; available at <http://warwick.ac.uk/snpm>) with 10 000 permutations (see Holmes et
205 al., 1996; Nichols & Holmes, 2001). This was computed separately by gender and the
206 resulting voxels were assessed for significance at 5% level and FWE-corrected, as
207 determined by permutation distribution. Similarly, to assess whether classification brain maps
208 significantly differ between genders in response to vocal/non-vocal sounds, this permutation
209 approach was implemented between groups (female > male, male > female) with 10 000
210 permutations and the resulting voxels were assessed for significance at 5% level and FWE-
211 corrected, as determined by permutation distribution (see Holmes et al., 1996; Nichols &
212 Holmes, 2001).

213 The between-group analysis was restricted to a mask defined by voxels with classification
214 accuracy significantly above theoretical chance ($p < 0.01$ uncorrected) in both females and
215 males. The resulting mask included 3783 voxels (yellow area in Fig. 2). The same mask was
216 applied for both, the univariate analysis and MVPA.

217 Separate brain maps of vocal vs. non-vocal contrast in female and male participants as well
218 as brain maps of contrasts between genders for both, univariate analysis and MVPA were
219 generated using the program MRICoGL (available at
220 <http://www.mccauslandcenter.sc.edu/mrico/mricron/>).

221 **Results**

222 **Univariate analysis: Vocal vs. non-vocal sounds**

223 The univariate analysis comparing activation to vocal and non-vocal sounds showed extended
224 areas of greater response to vocal sounds in the typical regions of the temporal voice areas

225 (TVA), highly similar for male and female subjects (Fig 1A). These regions were located
 226 bilaterally in the temporal lobes extending from posterior parts of the superior temporal
 227 sulcus (STS) along the superior temporal gyrus (STG) to anterior parts of the STS and also
 228 including several parts of the superior and middle temporal gyrus (STG, MTG).

229 Please insert Figure 1 here

230

231 Several hemispheric maxima of vocal vs. non-vocal response were located bilaterally along
 232 the superior temporal sulcus (STS) in both females and males (Fig. 1, Table 1). Fig.1A shows
 233 parameter estimates of the vocal > non-vocal contrasts at the maxima of the largest cluster
 234 sizes with the highest T-values of each hemisphere. The brain activation differences between
 235 vocal and non-vocal response was consistent across maxima in females (MNI coordinates
 236 left: x = -57, y= -16, z = -2, cluster size 3923, T = 20.85; right: x = 60, y= -13, z = -2, T-
 237 value = 20.64) and in males (MNI coordinates left: x = -60, y = -22, z = 1, cluster size 796,
 238 T=18.19; right: x = 60, y = -10, z = -2, cluster size 812, T-value = 17.46). Female listeners
 239 showed one large cluster covering the temporal lobes and subcortical parts of the brain. By
 240 contrast male listeners showed two separate voxel clusters in the left and right temporal lobes
 241 and no subcortical cluster connecting the two hemispheres (Table 1). Small bilateral clusters
 242 were found in inferior prefrontal cortex (inferior frontal gyrus, IFG) in both female and male
 243 listeners (p<0.05 FWE-corrected; Fig.1A).

244

245 Table 1. Voice-sensitive peak voxels of female and male RFX analysis (Univariate)

Anatomical location	Peak voxel			T values	Cluster size
	x	y	z		
Female					
Left/Right hemisphere					
Left STG, middle	-57	-16	-2	20.85	3923
Right STG, middle	60	-13	-2	20.64	
Right STG, middle	63	-22	-2	20.11	
Left frontal hemisphere					
IFG (pars triangularis)	-48	17	22	9.25	178
IFG (pars triangularis)	-39	29	-2	8.79	103
Precentral gyrus	-48	-7	43	6.32	5
Right frontal hemisphere					
IFG (orbital)	48	17	-8	4.87	1
Male					
Left hemisphere					
STG, middle	-60	-22	1	18.15	796
STG, middle	-57	-13	-2	17.97	
STG, posterior	-60	-37	4	12.73	
IFG (pars triangularis)	-42	29	-2	7.96	40
IFG (pars triangularis)	-42	17	22	5.56	32
Hippocampus	-18	-10	-14	4.61	1
Right hemisphere					
STG, middle	60	-10	-2	17.40	812
STG, middle	63	-22	-2	17.11	
STG, anterior	54	5	-14	11.51	
IFG (pars triangularis)	42	32	-2	6.76	165

IFG (pars triangularis)	54	23	22	6.62	
IFG (pars triangularis)	45	17	22	6.51	
Precentral gyrus	51	-1	46	7.60	22

Peak voxel coordinates in standard MNI space and corresponding t-values above for female and male 4.49 (FWE corrected p<0.05).

246

247 **MVPA analysis: vocal/non-vocal classification**

248 The MVPA analysis showed clusters of significantly above-chance voice/non-voice
 249 classification accuracy in TVA revealed by the univariate method above (Fig 1A, Table 2).
 250 Hemispheric maxima of classification accuracy were at comparable locations as the peaks of
 251 voice > non-voice activation revealed by the univariate method. The classification accuracy
 252 within the peak voxel of female listeners (MNI coordinates left: x = -60, y = -16, z = 1,
 253 cluster size 1676, T-value = 20.41; right: x = 66, y = -31, z = 4, cluster size 1671, T-value =
 254 21.45) as well as for male listeners (MNI coordinates left: x = -60, y = -22, z = 4, cluster size
 255 984, T-value = 13.70; right: x = 63, y = -28, z = 4, cluster size 1211, T-value = 16.07) were
 256 distinctly above the theoretical chance level of 0.5 (Fig.1B). Overall, the maximal
 257 classification accuracy was higher in female listeners as compared to male listeners at the
 258 peak voxels (Fig.1B, mean \pm s.e.m.: left peak in females 0.84 ± 0.006 , males 0.83 ± 0.009 ;
 259 right peak in females 0.85 ± 0.007 , males 0.84 ± 0.009 . Left peak in males 0.83 ± 0.009 ,
 260 females 0.85 ± 0.006 , right peak in males 0.85 ± 0.009 , females 0.87 ± 0.007). **Comparing**
 261 **MVPA and univariate analysis in Figure 1A and B, the MVPA analysis revealed more**
 262 **superficial cortical regions bilateral at the temporal pole, whereas the voxel cluster of the**
 263 **vocal vs. non-vocal difference of the univariate analysis extend more towards the midline of**
 264 **the brain.**

265 Table 2. Voice-sensitive peak voxels of female and male group analysis (MVPA)

Anatomical location	Peak voxel			T values	Cluster size
	x	y	z		
Female listeners					
Left hemisphere					
MTG, anterior	-60	-16	1	20.41	1676
MTG, posterior	-63	-37	7	18.26	
Right hemisphere					
STG, middle	66	-31	4	21.45	1671
STG, anterior	60	-7	-5	19.49	
Male listeners					
Left hemisphere					
MTG, middle	-60	-22	4	13.70	984
Right hemisphere					
STG, middle	63	-28	4	16.07	1211
MTG, anterior	63	-10	-5	14.88	

Peak voxel coordinates in standard MNI space and corresponding t-values above for female 4.38 and male 4.29 (FWE corrected p<0.05, as determined by permutation distribution with 10 000 permutations).

266 **Female vs. male contrasts**

267 The contrast of activation maps (univariate analysis) or classification accuracy maps
 268 (multivariate approach) from males and females revealed no significant voxels with greater
 269 parameter estimates for males > females at the chosen statistical significance threshold
 270 ($p < 0.05$, FWE-corrected) for either analysis methods. The reverse contrast (female > male),
 271 however, revealed significant voxel clusters showing greater parameter estimates for
 272 univariate analysis and higher classification accuracy for MVPA in female participants (Fig.
 273 2).

274 Table 3. Peak voxels of female > male contrast for univariate analysis and MVPA

Anatomical location	Peak voxel			T values	Cluster size	Cohen's d
	x	y	z			at the peak voxel
Univariate (female > male)						
Left hemisphere						
STG, posterior	-48	-34	16	4.02	4	0.48
Right hemisphere						
Insula	48	2	-5	4.04	1	0.49
MVPA (female > male)						
Left hemisphere						
STG, middle	-69	-19	-8	5.22	84	0.35
STG, middle	-66	-1	-8	5.02		
STG, middle	-51	-22	13	5.19	156	0.35
STG, middle	-48	-31	4	4.77		
STG, posterior	-42	-43	7	4.66		
MTG, middle	-57	-55	16	3.82	2	
MTG, middle	-69	-40	1	3.80	2	
STG, middle	-69	-10	10	3.79	2	
Right hemisphere						
STG, middle	69	-7	-11	4.48	52	0.24
MTG, middle	66	-22	-11	4.42		
MTG, middle	69	-34	1	3.70	1	

Peak voxel coordinates in standard MNI space and corresponding t-values above 3.85 (univariate analysis, FWE corrected $p < 0.05$) and 3.70 for MVPA (FWE corrected $p < 0.05$, as determined by permutation distribution with 10 000 permutations) and Cohen's d for large cluster size. The Cohen's d of the MVPA refers to the mean difference in classification accuracy (contrast estimates of the univariate analysis respectively), divided by the pooled standard deviation for those means.

275

276 When analysed with the univariate approach (Fig. 2A) the contrast female > male yielded
 277 only a few significant voxels: One cluster consisted of four voxels in the left posterior part of
 278 STG and only one voxels in the right Insula (Fig. 2A, Table 3). The corresponding contrast
 279 estimates for the reported peak voxels (MNI coordinates left: $x = -48$, $y = -34$, $z = 16$, cluster
 280 size 4, $T = \text{value} = 4.02$; right: $x = 48$, $y = 2$, $z = -5$, cluster size = 1, $T\text{-value} = 4.04$) showed a
 281 positive response for females in both hemispheres and for the left hemisphere in males. The
 282 Cohen's d effect size values ($d = 0.48$ and 0.49) suggested a moderate difference at the peak

283 voxel (Table 3). Overall, females showed a stronger activation in response to vocal vs. non-
284 vocal sounds as compared to males at both maxima (Fig. 2A).

285

286 Please insert Figure 2 here

287

288 The female > male contrast of classification accuracy maps identified significant voxel
289 clusters in the middle part of the middle temporal gyrus (MTG) in both hemispheres, in
290 which classification accuracy was greater for female than male subjects (red clusters in Fig.
291 2B). Areas of greater classification accuracy in females were more extended in the left
292 hemisphere with an additional smaller cluster located in the superior temporal gyrus (STG).
293 The peak voxels of female > male classification accuracy difference were located in the
294 middle part of the MTG (bilateral), and the left middle STG (MNI coordinates left: $x = -69$, $y = -19$, $z = -8$, cluster size 84, T-value = 5.22; $x = -51$, $y = -22$, $z = 13$, cluster size 156, T-
295 value = 5.19; right: $x = 69$, $y = -7$, $z = -11$, cluster size 52, T-value = 4.48; cf. circle in Fig
296 2A). The Cohen's d effect size values ($d = 0.35$, 0.35 and 0.24) suggested a small difference
297 at the peak voxel (Table 3). Classification accuracy (computed in native space) at these
298 coordinates was distinctly above chance (50%) for both females and males, but higher in
299 females across peaks (Fig. 2B).

301 Discussion

302 The present study aimed to investigate gender differences on voice localizer scans by
303 employing the conventional univariate analysis as well as MVPA. Both analysis approaches
304 revealed largely overlapping/comparable and robust estimates of the TVAs in female and
305 male listeners. However, the MVPA was more sensitive to differences in the middle MTG of
306 the left and right hemispheres and the middle left STG between genders as compared to
307 univariate analysis with higher classification accuracy in women.

308

309 Robust TVAs

310 The estimated TVAs using MVPA robustly replicated and confirmed prior fMRI findings
311 applying the voice localizer (Belin et al., 2002, 2000; Belin & Zatorre, 2003; Scott &
312 Johnsrude, 2003; Von Kriegstein et al., 2003). Both analysis methods showed comparable
313 maps of classification accuracy (MVPA) and of vocal vs. non-vocal activity difference
314 (univariate analysis) for both female and male listeners. The average classification accuracy
315 at the peak voxel was distinctly above chance level and higher in female as compared to male
316 listeners. The peak voxels were at comparable locations (along middle and posterior parts of
317 the STS) for both analysis approaches and both genders. A small difference between the
318 MVPA and univariate analysis can be seen bilateral at the temporal pole, where the MVPA
319 detected more vocal/non-vocal differences in superficial cortical regions as compared to the
320 univariate analysis. In addition to the activation brain maps showing the robustly estimated
321 TVAs (univariate analysis), the MVPA results extend previous findings by providing a
322 corresponding classification accuracy brain map. When brain maps are considered for each
323 analysis approach and for female and male listeners separately, our findings showed no
324 distinct differences between genders and between univariate analysis and MVPA. Instead
325 comparable voxel clusters of a similar size in the bilateral temporal lobes were identified,
326 verifying the prior univariate analysis and the robustness of the TVAs (see e.g. Belin et al.
327 2000).

328

329 **Gender differences**

330 When data were analysed with MVPA, differences between female and male listeners in
331 response to vocal / non-vocal sounds were found by contrasting female > male (but not male
332 > female). A significant difference in success of the MVPA between female and male
333 listeners was apparent in the middle part of the MTG in both hemispheres and in the middle
334 part of the STG in the left hemisphere. Effect sizes showed a small difference at the peak
335 voxels. Despite the large sample size used in this study, the univariate analysis showed no
336 major activation differences between genders. Only two small clusters with one to four
337 voxels were significant in the posterior and anterior part of the STG. In the univariate
338 analysis, the overall activation difference between vocal vs. non-vocal sounds was stronger in
339 female as compared to male listeners and effect sizes showed a moderate difference at the
340 peak voxels.

341 The distinct gender differences located in the middle part of MTG and middle part of STG
342 between genders revealed by the MVPA survived our applied criteria (FWE-correction). In
343 these regions, the classifier successfully distinguished between the vocal and non-vocal
344 condition with better overall accuracy in females as compared to males across the peak
345 voxels. Thus, BOLD signal in parts of auditory cortex seem to carry less information for
346 discriminating vocal from nonvocal sounds in male than females listeners. We do not make
347 any inference on the nature of the underlying processing differences in terms of mental states
348 or cognitive mechanisms, but possible explanations for our findings are discussed below.

349 MVPA may overall be more sensitive to detect small differences in the activation patterns to
350 vocal and non-vocal sounds. Thus, differences between genders appear significant only when
351 analysed with MVPA (Haynes et al., 2007; Kriegeskorte et al., 2006; Norman et al., 2006).
352 The differences in classification accuracy between female and male listeners, identified in
353 parts of auditory cortex, may be contributed to by a different predisposition of female/male
354 listeners to the presented vocal sound samples of the voice localizer. Previous findings
355 suggest a sex-difference in response to infant crying and laughing. Women showed a
356 deactivation in the anterior cingulate cortex (ACC) to both laughing and crying (independent
357 of parental status) as compared to men (Seifritz et al., 2003). In contrast, another study
358 showed increased activation to infant vocalization in the amygdala and ACC whereas men
359 showed increased activation to the control stimuli (fragment recombined and edge smoothed
360 stimuli of the original laughing/crying samples). This may reflect a tendency in women for a
361 response preference to infant vocal expressions (Sander et al., 2007). A recent study by De
362 Pisapra et al. in 2013 found a sex-difference in response to a baby cry. Women decreased
363 brain activity in DPFC regions and posterior cingulate cortex when they suddenly and
364 passively heard infant cries, whereas men did not. They interpreted their findings in such a
365 way that the female brain interrupts on-going mind-wandering during cries and the male
366 brain continues in self-reflection (De Pisapia et al., 2013). In our study half of the vocal
367 stimuli consisted of infant vocalizations (also emotional expressions such as laughing and
368 crying) and our results may reflect differences in the fine-grained pattern of distributed
369 activity in female and male listeners in response to these vocal expressions of children and
370 babies. The outcome in this study may be affected by anatomical differences in brain
371 structure/size between female and male listeners (Brett et al., 2002). In general individuals
372 vary in their anatomical brain structures and undergo the experiment with different mental
373 states which may influence their brain responses (Huettel et al., 2008).

374 To date, there is also evidence for differences in the vocal processing and in particular in
375 speech perception between genders from both behavioral (J. A. Hall, 1978, Skuk and
376 Schweinberger, 2013) and previous fMRI studies (Junger et al., 2013; Schirmer et al., 2002,

377 2007, 2004; Shaywitz et al., 1995). These studies found activation differences in frontal brain
378 regions (Junger et al., 2013, Schirmer et al.,2004) and the left posterior MTG and the angular
379 gyrus (Junger et. al.,2004). The deviation of the current results in terms of identified brain
380 regions may be due to the different experimental design and computed contrasts, the different
381 applied criteria (e.g. mask), number of included participants and implemented analysis
382 methods. Future studies should further aim to elucidate the relationships between behavioral
383 and functional activation differences. However, the current study shows that the choice of
384 fMRI analysis method (e.g. MVPA) is of relevance when considering subtle between-gender
385 differences.

386 Regarding the current study, it would be interesting to separate the different vocal categories
387 in the analysis (e.g. by speaker: female/male adults vs. infants/babies) and to perform a
388 behavioral task in order to link differences in brain activation to behavior of the listener.
389 Furthermore, it would be interesting for future studies to take into account more specific
390 aspects of voice quality, which were not considered in the current study. Even subtle
391 differences in phonation (e.g. whispery voice, harshness of a voice), articulation (e.g. vowel
392 space) and or prosody (e.g. pitch variability, loudness, tempo) are critical aspects of voice
393 processing and could be investigated using similar methodical approaches. Apart from
394 studying differences between women and men, also other listener characteristics, such as
395 differences between young and elderly participants, different nationalities and/or familiarity
396 with the presented voices/stimuli should be considered.

397

398 **Conclusion**

399 Male and female participants were similar in their pattern of activity differences in response
400 to vocal vs. nonvocal sounds in the temporal voice areas of the auditory cortex. Yet, MVPA
401 revealed several regions of significant gender differences in classification performance
402 between female and male listeners: in these regions the distributed pattern of local activity
403 from female participants allowed significantly better vocal/nonvocal classification than that
404 of male participants; no region showed the opposite male > female difference. The neuronal
405 mechanisms underlying the observed differences remain unclear.

406 **Acknowledgement**

407 We thank David Fleming for helpful matlab support and Chris Benwell for English
408 corrections.

409 **References**

410 Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice
411 Perception. *British Journal of Psychology*, *102*, 711–725.
412 doi:http://dx.doi.org/10.1111/j.2044-8295.2011.02041.x

413 Belin, P., Shirley, F., & Catherine, B. (2004). Thinking the voice: neural correlates of voice
414 perception. *Trends in Cognitive Science*, *8*(3), 129–135.
415 doi:http://dx.doi.org/10.1016/j.tics.2004.01.008

416 Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal
417 lobe. *Neuroreport*, *14*(16), 2105–9. doi:http://dx.doi.org/10.1097/00001756-200311140-
418 00019

- 419 Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds.
420 *Cognitive Brain Research*, *13*(1), 17–26. doi:<http://dx.doi.org/10.1016/S0926->
421 [6410\(01\)00084-2](http://dx.doi.org/10.1016/S0926-6410(01)00084-2)
- 422 Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in
423 human auditory cortex. *Nature*, *403*(6767), 309–12. doi:[10.1038/35002078](http://dx.doi.org/10.1038/35002078)
- 424 Binder, J. R., Frost, J. a, Hammeke, T. a, Bellgowan, P. S., Springer, J. a, Kaufman, J. N., &
425 Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds.
426 *Cerebral cortex*, *10*(5), 512–28. doi:<http://dx.doi.org/10.1093/cercor/10.5.512>
- 427 Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in
428 the human brain. *Nature reviews. Neuroscience*, *3*(3), 243–9. doi:[10.1038/nrn756](http://dx.doi.org/10.1038/nrn756)
- 429 Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain
430 reading”: detecting and classifying distributed patterns of fMRI activity in human visual
431 cortex. *NeuroImage*, *19*(2), 261–270. doi:<http://dx.doi.org/10.1016/S1053->
432 [8119\(03\)00049-1](http://dx.doi.org/10.1016/S1053-8119(03)00049-1)
- 433 De Pisapia, N., Bornstein, M. H., Rigo, P., Esposito, G., De Falco, S., & Venuti, P. (2013).
434 Sex differences in directional brain responses to infant hunger cries. *NeuroReport*, *24*(3),
435 142–146. doi:[10.1097/WNR.0b013e32835df4fa](http://dx.doi.org/10.1097/WNR.0b013e32835df4fa)
- 436 Ethofer, T., Bartscher, J., Geschwind, M., Benjamin, K., Wildgruber, D., & Vuilleumier, P.
437 (2012). Emotional Voice Areas: Anatomic Location, Functional Properties, and
438 Structural Connections Revealed by Combined fMRI/DTI. *Cerebral Cortex*, *22*(1), 191–
439 200. doi:<http://dx.doi.org/10.1093/cercor/bhr113>
- 440 Hall, G. B. C., Witelson, C. A. S. F., Szechtman, H., & Nahmias, C. (2004). Sex difference in
441 functional activation patterns revealed by increased emotion processing demands, *15*(2).
442 doi:[10.1097/01.wnr.0000101310.64109.94](http://dx.doi.org/10.1097/01.wnr.0000101310.64109.94)
- 443 Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*(4),
444 845–857. doi:[10.1037/0033-2909.85.4.845](http://dx.doi.org/10.1037/0033-2909.85.4.845)
- 445 Hall, J. A., Murphy, N. A., & Mast, M. S. (2006). Recall of Nonverbal Cues: Exploring a
446 New Definition of Interpersonal Sensitivity. *Journal of Nonverbal Behavior*, *30*(4), 141–
447 155. doi:[10.1007/s10919-006-0013-3](http://dx.doi.org/10.1007/s10919-006-0013-3)
- 448 Haynes, J.-D., & Rees, G. (2005). Predicting the stream of consciousness from activity in
449 human visual cortex. *Current biology : CB*, *15*(14), 1301–7.
450 doi:<http://dx.doi.org/10.1016/j.cub.2005.06.026>
- 451 Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans.
452 *Nature reviews. Neuroscience*, *7*(7), 523–34. doi:<http://dx.doi.org/10.1038/nrn1931>
- 453 Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007).
454 Reading hidden intentions in the human brain. *Current biology : CB*, *17*(4), 323–8.
455 doi:<http://dx.doi.org/10.1016/j.cub.2006.11.072>

- 456 Holmes, A. P., Blair, R. C., Watson, J. D., & Ford, I. (1996). Nonparametric analysis of
457 statistic images from functional mapping experiments. *Journal of cerebral blood flow*
458 *and metabolism*, *16*(1), 7–22. doi:10.1097/00004647-199601000-00002
- 459 Huettel, S. A., Song, A. W., & McCarthy, G. (2008). *Functional Magnetic Resonance*
460 *Imaging* (2nd ed., p. 510). Massachusetts: Sinauer Associates.
- 461 Junger, J., Pauly, K., Bröhr, S., Birkholz, P., Neuschaefer-Rube, C., Kohler, C., ... Habel, U.
462 (2013). Sex matters: Neural correlates of voice gender perception. *NeuroImage*, *79*,
463 275–87. doi:http://dx.doi.org/10.1016/j.neuroimage.2013.04.105
- 464 Kotz, S. A., Kalberlah, C., Bahlmann, J., Friederici, A. D., & Haynes, J.-D. (2013).
465 Predicting vocal emotion expressions from the human brain. *Human brain mapping*,
466 *34*(8), 1971. doi:http://dx.doi.org/10.1002/hbm.22041
- 467 Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., & Wildgruber, D. (2009). Cerebral
468 representation of non-verbal emotional perception: fMRI reveals audiovisual integration
469 area between voice- and face-sensitive regions in the superior temporal sulcus.
470 *Neuropsychologia*, *47*(14), 3059–66.
471 doi:http://dx.doi.org/10.1016/j.neuropsychologia.2009.07.001
- 472 Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain
473 mapping. *Proceedings of the National Academy of Sciences of the United States of*
474 *America*, *103*(10), 3863–8. doi:http://dx.doi.org/10.1073/pnas.0600244103
- 475 Latinus, M., Crabbe, F., & Belin, P. (2011). Learning-induced changes in the cerebral
476 processing of voice identity. *Cerebral cortex*, *21*(12), 2820–8.
477 doi:http://dx.doi.org/10.1093/cercor/bhr077
- 478 Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with
479 pattern-information fMRI: an introductory guide. *Scan*, 1–9.
- 480 Nichols, T. E., & Holmes, A. P. (2001). Nonparametric permutation tests for functional
481 neuroimaging: a primer with examples. *Human brain mapping*, *15*(1), 1–25.
- 482 Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading:
483 multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424–30.
484 doi:10.1016/j.tics.2006.07.005
- 485 Sander, K., Frome, Y., & Scheich, H. (2007). FMRI activations of amygdala, cingulate
486 cortex, and auditory cortex by infant laughing and crying. *Human Brain Mapping*,
487 *28*(10), 1007–1022. doi:10.1002/hbm.20333
- 488 Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms
489 mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30.
490 doi:10.1016/j.tics.2005.11.009
- 491 Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional
492 prosody during word processing. *Cognitive Brain Research*, *14*(2), 228–223.
493 doi:http://dx.doi.org/10.1016/S0926-6410(02)00108-8

- 494 Schirmer, A., Simpson, E., & Escoffier, N. (2007). Listen up! Processing of intensity change
495 differs for vocal and nonvocal sounds. *Brain research*, *1176*, 103–12.
496 doi:<http://dx.doi.org/10.1016/j.brainres.2007.08.008>
- 497 Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive
498 processing of vocal emotional expressions. *Neuroreport*, *16*(6), 635–9. Retrieved from
499 <http://www.ncbi.nlm.nih.gov/pubmed/15812323>
- 500 Schirmer, A., Zysset, S., Kotz, S. A., & Yves von Cramon, D. (2004). Gender differences in
501 the activation of inferior frontal cortex during emotional speech perception.
502 *NeuroImage*, *21*(3), 1114–1123. doi:<http://dx.doi.org/10.1016/j.neuroimage.2003.10.048>
- 503 Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of
504 speech perception. *Trends in neurosciences*, *26*(2), 100–7.
505 doi:[http://dx.doi.org/10.1016/S0166-2236\(02\)00037-1](http://dx.doi.org/10.1016/S0166-2236(02)00037-1)
- 506 Seifritz, E., Esposito, F., Neuhoff, J. G., Lüthi, A., Mustovic, H., Dammann, G., ... Di Salle,
507 F. (2003). Differential sex-independent amygdala response to infant crying and laughing
508 in parents versus nonparents. *Biological Psychiatry*, *54*(12), 1367–1375.
509 doi:[10.1016/S0006-3223\(03\)00697-8](http://dx.doi.org/10.1016/S0006-3223(03)00697-8)
- 510 Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R.
511 K., ... Katz, L. (1995). Sex differences in the functional organization of the brain for
512 language. *Nature*, *373*(6515), 607–9. doi:[10.1038/373607a0](http://dx.doi.org/10.1038/373607a0)
- 513 Skuk, V. G., & Schweinberger, S. R. (2013). Gender differences in familiar voice
514 identification. *Hearing research*, *296*, 131–40.
515 doi:<http://dx.doi.org/10.1016/j.heares.2012.11.004>
- 516 Von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural
517 responses to speech by directing attention to voices or verbal content. *Cognitive brain*
518 *research*, *17*(1), 48–55. doi:[http://dx.doi.org/10.1016/S0926-6410\(03\)00079-X](http://dx.doi.org/10.1016/S0926-6410(03)00079-X)
- 519

520 **Figure legends**

521 Figure 1: Brain maps of female (red, n = 149) and male (blue, n = 123) participants. (A)
522 Univariate analysis showing bilateral activation along the superior temporal sulcus (STS) and
523 in the inferior frontal gyrus (IFG) and corresponding contrast estimates of vocal vs. non-vocal
524 sounds plotted for peak voxel (one-sample t-test, FWE-corrected, $p < 0.05$; cf. circles, note
525 that the two peaks with highest T-value and largest cluster size are indicated per group). (B)
526 MVPA showing comparable classification accuracy maps along STS, but not IFG and
527 average classification accuracy \pm S.E.M. at peak voxel (calculated in native space) was
528 distinctly above chance level (0.5) for both females and males (maximum intensity projection
529 of *t*-statistic image threshold at FWE-corrected $p < 0.05$, as determined by permutation
530 distribution with 10 000 permutations).

531 Figure 2: Contrast between female > male (red). (A) Univariate analysis showing significant
532 female > male difference (two-sample t-test, FWE-corrected, $p < 0.05$) in the left posterior
533 part of the superior temporal gyrus (STG) and the right anterior STG. Contrast estimates at
534 peak voxel showing stronger activation in females (black) as compared to males (gray) in
535 response to vocal vs. non-vocal sounds. (B) MVPA showing significant classification
536 accuracy above chance level in the right middle part of the middle temporal gyrus and the
537 right middle STG as well as in the left middle MTG with higher average classification
538 accuracy in females (black) than in males (gray) (maximum intensity projection of *t*-statistic
539 image threshold at FWE-corrected $p < 0.05$, as determined by permutation distribution with
540 10 000 permutations). The (yellow) cluster shows the mask including voxels with
541 significantly above chance classification accuracy in both females and males.

Figure 1.TIF

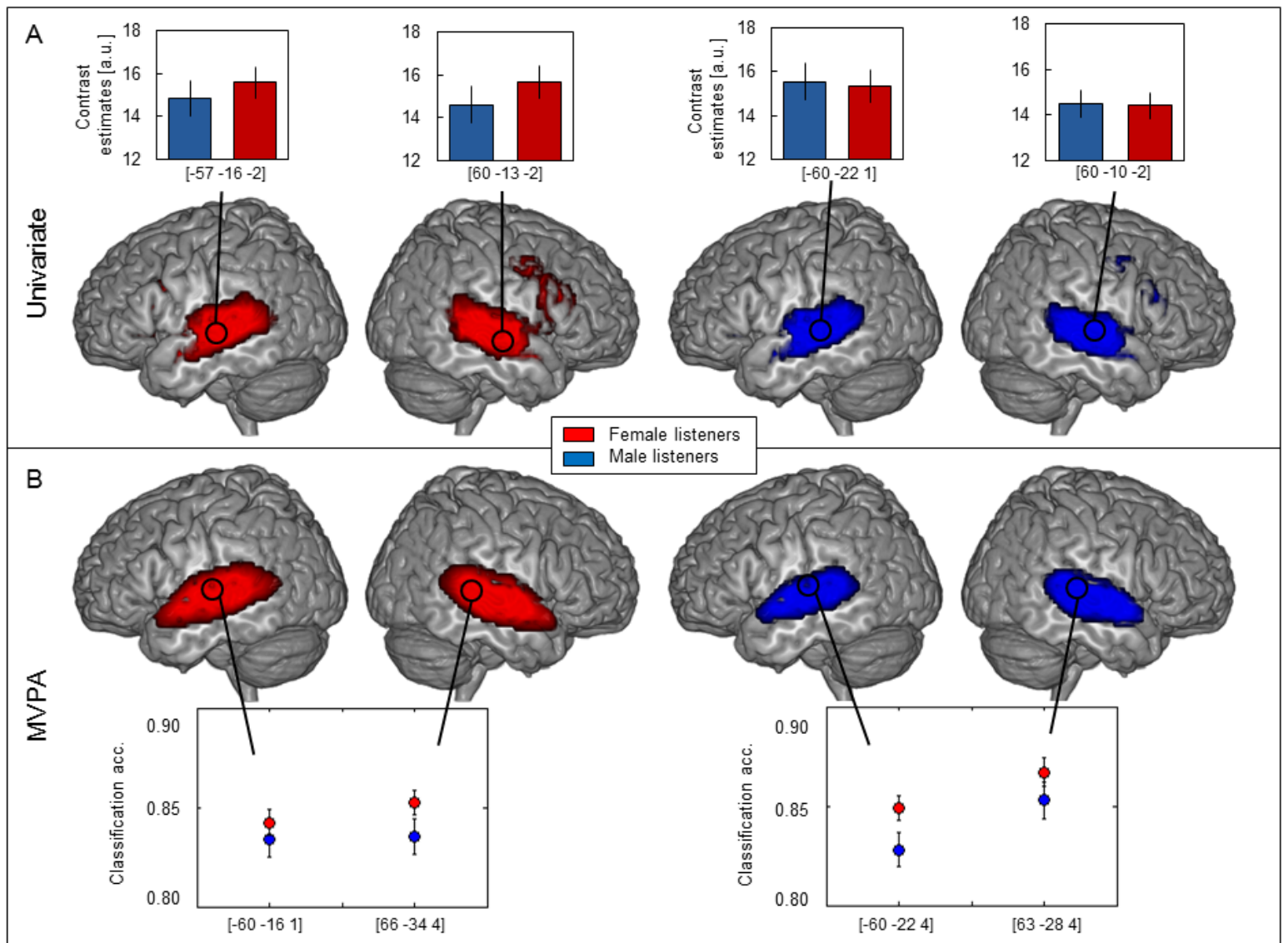


Figure 2.TIF

