doi:10.1093/scan/nsm028 SCAN (2007) 2, 334–337

The voices of seduction: cross-gender effects in processing of erotic prosody

Thomas Ethofer,[1,2] Sarah Wiethoff,[1,2] Silke Anders,[2] Benjamin Kreifelts,[1] Wolfgang Grodd,[2] and Dirk Wildgruber[1]

[1]Department of General Psychiatry and [2]Section Experimental MR of the CNS, Department of Neuroradiology, University of Tuebingen, Tuebingen, Germany

Gender specific differences in cognitive functions have been widely discussed. Considering social cognition such as emotion perception conveyed by non-verbal cues, generally a female advantage is assumed. In the present study, however, we revealed a cross-gender interaction with increasing responses to the voice of opposite sex in male and female subjects. This effect was confined to erotic tone of speech in behavioural data and haemodynamic responses within voice sensitive brain areas (right middle superior temporal gyrus). The observed response pattern, thus, indicates a particular sensitivity to emotional voices that have a high behavioural relevance for the listener.

Keywords: emotion; erotic; fMRI; prosody; sex; voice

INTRODUCTION

Emotional significance of a stimulus or event is tied to its potential to further or obstruct a person's goals (Ellsworth and Scherer, 2003). Thus, emotions reflect a meaning-centred system which is based on appraisal processes that evaluate the behavioural relevance of events for the organism and as such they have a greater flexibility than stimulus-centred systems such as reflexes (Smith and Lazarus, 1990). Automatic appraisal of emotional information is highly important for an individual's well-being since it is inherently necessary for avoidance of danger and successful social interactions, such as forming friendships and finding mating partners. In the auditory modality, emotional information can be expressed by modulation of speech melody (prosody). Recently, enhanced responses in voice-processing areas (Belin *et al.*, 2000; von Kriegstein and Giraud, 2004; von Kriegstein *et al.*, 2005; Warren *et al.*, 2006) to angry (Grandjean *et al.*, 2005; Ethofer *et al.*, 2006) and happy (Ethofer *et al.*, 2006) relative to neutral prosody have been demonstrated. However, any comparison between brain responses to different prosodic categories is influenced by differences in low-level acoustic parameters (Banse and Scherer, 1996). Are differences in acoustic properties sufficient to explain stronger responses of voice processing modules to emotional relative to neutral prosody or is neuronal activity of this region additionally modulated by the behavioural relevance of the stimuli? Erotic prosody offers an opportunity to clarify this question since its behavioural relevance (i.e. the prospect of a potential sexual partner) is dependent on the gender of speaker and listener. Critically, such cross-gender interactions are independent from possible differences in stimulus properties since they are based on between-subject comparisons to the same physical stimuli.

Received 29 March 2007; Accepted 25 May 2007

Advance Access publication 29 June 2007

This study was supported by the 'Junior science program of the Heidelberger Academy of Sciences and Humanities' and the German Research Foundation (SFB 550 B10).

Correspondence should be addressed to Dirk Wildgruber, Department of General Psychiatry, University of Tuebingen, Osianderstrasse 24, 72076 Tuebingen, Germany. E-mail: dirk.wildgruber@med.uni-tuebingen.de

MATERIAL AND METHODS

Stimuli and task

The stimulus set employed in the fMRI experiment comprised 25 nouns and 25 adjectives the semantic content of which was previously rated (Herbert *et al.*, 2006) as emotionally neutral and low-arousing on a 9-point self assessment manikin (SAM) scale (Bradley and Lang, 1994). These stimuli were selected from a pool of 126 words spoken by six professional actors (three female, three male) in five different speech melodies corresponding to the prosodic categories neutral, anger, fear, happiness and eroticism. All stimuli were normalized to the same peak intensity and balanced for gender of the speaker, number of syllables and word frequency over prosodic categories. A pre-study (10 subjects, mean age 24 years, 4 female, 6 male) was conducted to ensure that the prosodic category of all stimuli employed in the main experiment is correctly identified by at least 70% of the subjects.

The fMRI experiment consisted of two sessions. In both sessions, the same 50 stimuli were presented binaurally via magnetic resonance-compatible headphones with piezo-electric signal transmission (Jaencke *et al.*, 2002) in a passive listening paradigm. The order of stimulus presentation was fully randomized. Stimulus onset was jittered relative to scan onset in steps of 500 ms and the inter stimulus interval ranged from 9–12 s.

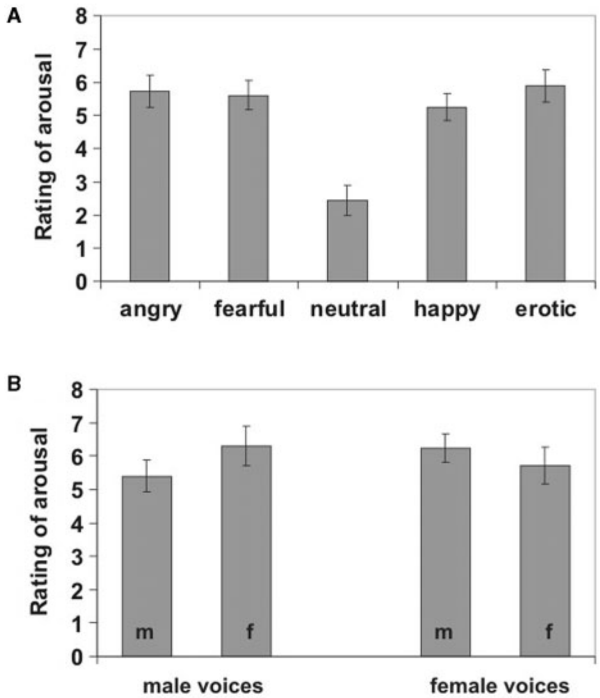

Fig. 1 (A) Rating of emotional arousal of the five prosodic categories on a 9-point scale ranging from 0 (totally unexciting) to 8 (extremely exciting). (B) Rating of emotional arousal of erotic prosody dependent on the gender of speaker and listener (m = male listener, f = female listener). Error bars represent standard errors of the mean.

MRI acquisition

MR images were acquired using a 3 T Siemens TRIO scanner. After obtaining a static field-map for off-line image distortion correction of echoplanar imaging (EPI) scans, two series of 300 EPI scans (25 slices, 64×64 matrix, TR = 1.83 s, TE = 40 ms) were acquired. Subsequently, a magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence was employed to acquire high-resolution ($1 \times 1 \times 1\,mm^3$) T_1-weighted structural images (TR = 2300 ms, TE = 3.93 ms, TI = 1100 ms).

Data analysis

Functional images were analyzed using statistical parametrical mapping software (SPM2, Wellcome Department of Imaging Neuroscience, London, UK). Pre-processing comprised realignment to the first volume of the time series, correction of image distortions by use of a static field map (Andersson *et al.*, 2001), normalization into MNI (Montreal Neurological Institute) space (Collins *et al.*, 1994) and spatial smoothing with an isotropic Gaussian filter (10 mm full width at half maximum). Statistical analysis relied on a general linear model (Friston *et al.*, 1994) and a random effects analysis was performed to investigate which brain regions respond stronger to emotional than to neutral intonations. Activations are reported at a height threshold of $P < 0.001$ and significance was assessed at the cluster level with an extent threshold of $P < 0.05$ (corresponding to a minimal cluster size of 50 voxels), corrected for multiple comparisons across the whole brain. To investigate the influence of cross-gender interactions at behavioural and neurophysiologic level, arousal ratings obtained in the behavioural study and fMRI parameter estimates of the most significantly activated voxel in right mid STG as defined by the contrast (emotional > neutral prosody) were submitted to two-factorial repeated measures ANOVAs with gender of the speaker as within- and gender of the listener as between-subject factor. Paired *t*-tests were employed to determine whether these cross-gender interactions were significantly stronger for erotic prosody than for the other prosodic categories.

RESULTS AND DISCUSSION

Words spoken in erotic, happy, neutral, angry or fearful prosody were presented during a behavioural and an fMRI experiment. In the behavioural experiment, 20 healthy right-handed heterosexual subjects (mean age 25.2 years, 10 female, 10 male) rated the arousal of emotional prosody. Higher arousal ratings were obtained for all four emotional categories than for neutral prosody [all paired $T(19) > 6.5$, $P < 0.001$, Figure 1A]. A two-factorial repeated measures ANOVA with gender of the speaker as within- and gender of the listener as between-subject factor revealed no main effect [$F(1,18) < 1$], but a significant interaction [$F(1,18) = 16.77$, $P < 0.001$] on arousal ratings of erotic prosody. This interaction was attributable to higher arousal ratings of stimuli spoken by actors of opposite than same sex as the listener (Figure 1B). No significant interaction was found for any of the other four prosodic categories [all $F(1,18) < 1$]. Furthermore, the cross-gender interaction was significantly stronger for erotic than for the other prosodic categories [paired $T(23) = 2.78$, $P < 0.01$, one-tailed] demonstrating that this effect was not due to overall higher arousal ratings of voices of opposite relative to same sex as the listener, but occurred specifically for erotic prosody.

In the fMRI experiment, the same stimulus set was presented to a different group of subjects comprising 24 healthy right-handed heterosexual adults (mean age 25.1 years, 12 females, 12 males) in a passive-listening paradigm. Comparison of haemodynamic responses to emotionally and neutrally spoken words revealed a cluster in the right superior temporal gyrus with two distinct maxima, which were situated in the primary auditory cortex and in the associative auditory cortex of the mid STG (Figure 2A). Other brain regions showing stronger activations to emotional than to neutral intonations included the left temporal pole, hypothalamus and three small clusters within the left superior and middle temporal gyrus (Table 1). Separate parameter estimates obtained from the right mid STG for the prosodic categories demonstrate that stimuli of all four emotional categories elicited stronger responses than neutral stimuli [Figure 2C, angry *vs* neutral $T(23) > 2.63$, fearful *vs* neutral $T(23) = 2.71$, happy *vs* neutral

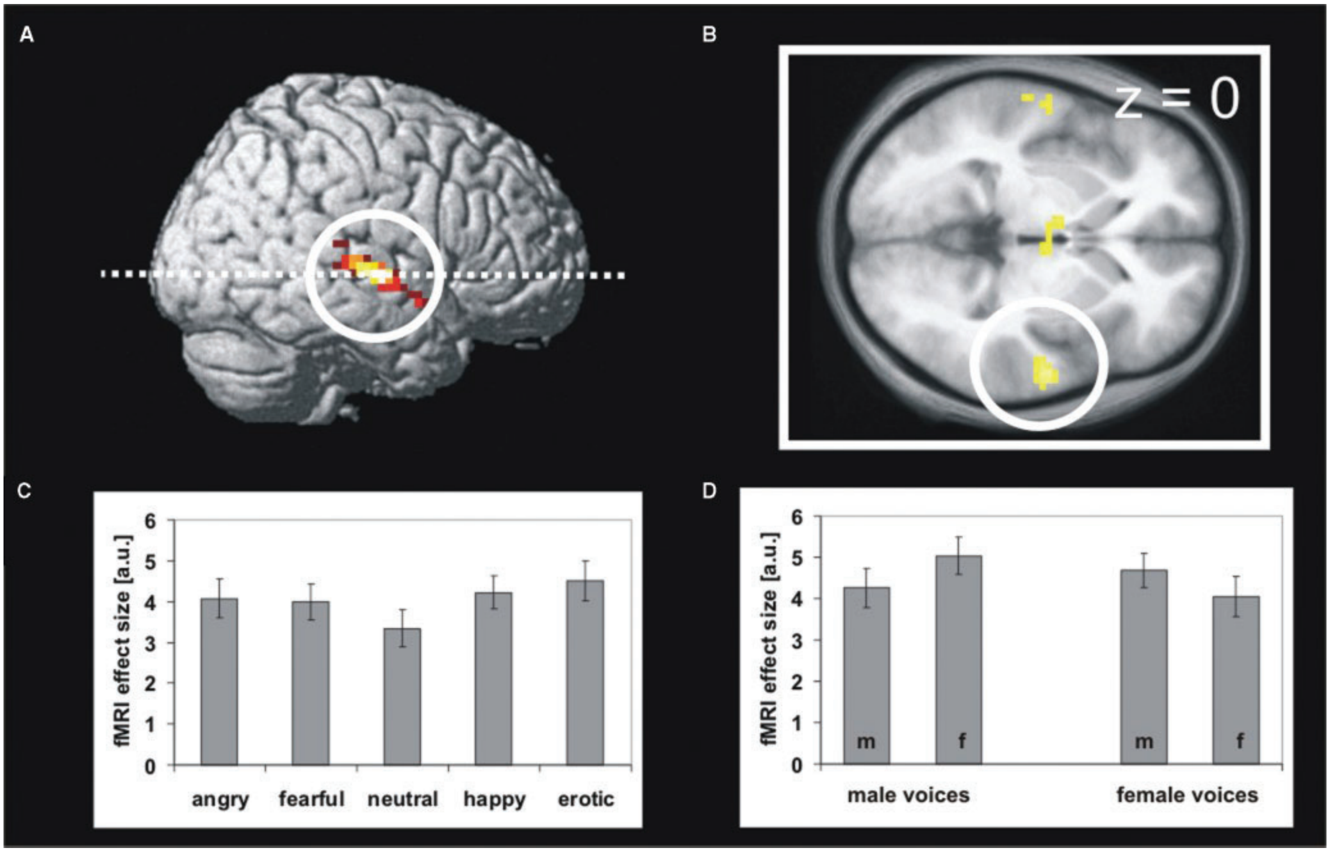

Fig. 2 (A) Brain regions showing stronger activations to emotional than neutral prosody rendered on a right hemisphere of a standard brain and (B) a transversal slice ($z = 0$) of the mean T_1 image obtained from the normalized brains of the fMRI participants. The activation cluster in the right mid STG is marked by a white circle. (C) Effect size of fMRI responses in right mid STG in arbitrary units (a.u.) for the five prosodic categories and (D) for erotic prosody dependent on gender of speaker and listener (m = male listener, f = female listener). Error bars represent standard errors of the mean.

Table 1 Activation during perception of emotional prosody (*vs* neutral tone of speech)

Anatomical definition	MNI coordinates			*Z*-score	Cluster size
Right auditory cortex					99*
Primary auditory cortex (Heschl's gyrus)	48	−24	3	4.06	
Associative auditory cortex (mid STG)	63	−12	0	3.72	
Left temporal pole	−36	6	−18	4.10	48
Hypothalamus	9	0	−6	3.74	44
Left superior temporal gyrus	−60	−9	3	3.26	7
Left middle temporal gyrus	−63	−21	0	3.30	6
Left middle temporal gyrus	−51	−36	6	3.25	6

*$P < 0.05$, corrected for multiple comparisons within the whole brain.

$T(23) = 4.41$, erotic *vs* neutral $T(23) = 4.86$, all $P < 0.01$, one-tailed]. These findings are in agreement with previous findings on processing of emotional information in the voice (Grandjean *et al.*, 2005; Ethofer *et al.*, 2006) and parallel results obtained for the visual domain (Surguladze *et al.*, 2003) that suggest prioritized processing of arousing stimuli in face sensitive regions for a broad spectrum of emotional categories. Parameter estimates extracted from right mid STG were submitted to a two-factorial ANOVA for repeated measures with gender of the speaker as within- and gender of the fMRI participant as between-subject factor. This analysis revealed no main effect [$F(1,22) < 1$], but a significant interaction [$F(1,22) = 5.7$, $P < 0.05$] on brain responses of right mid STG to erotic prosody. In analogy to the behavioural study, this interaction was due to stronger responses to voices of opposite than same sex (Figure 2D) and specific for erotic prosody (all $F < 1$ for the other four prosodic categories). Again, the interaction was significantly stronger for erotic than for the other four prosodic categories [$T(23) = 1.96$, $P < 0.05$, one-tailed].

Our data demonstrate that both men and women attribute higher arousal ratings to erotic prosody expressed by speakers of opposite than same sex and that this effect is mirrored by the response pattern of voice-processing cortices in right mid STG. The specificity of this effect for erotic prosody indicates that this region is not generally tuned to voices of opposite sex, but that such modulation depends on whether the signal is of higher behavioural relevance if spoken by a conspecific of opposite sex. Enhancement of subjective arousal ratings and mid STG responses to erotic prosody of opposite as compared to same sex was similar for male and female subjects. This finding contrasts with results obtained in the visual domain were the difference in arousal ratings and cortical reactivity to pictures of opposite sex relative to same sex nudes (Costa *et al.*, 2003) or erotic couples (Karama *et al.*, 2003; Sabatinelli *et al.*, 2004) is larger in men than in women. Future research should address the question whether gender-specific differences in brain activation of visual areas are due to the fact that previous studies used stimuli which are less arousing for women than for men or generally reflect a stronger visual orientation of men for sexual selection criteria. Furthermore, differentiation of brain responses to stimuli inducing sexual arousal from those to mating signals which are recognized as such, but cause other affective reactions because the recipient does not share the sender's interest, await further investigation. Decoding of mating signals in the voice is important for successful reproduction in many species ranging from invertebrates (Kiflawi and Gray, 2000) and amphibians (Boul *et al.*, 2007) to primates (Hauser, 1993). Thus, the results presented here could inspire new comparisons between species in the neuronal correlates underlying comprehension of auditory mating signals, a class of stimuli of high relevance for an individual's well-being and survival of its species.

Conflict of Interest

None declared.

REFERENCES

Andersson, J.L., Hutton, C., Ashburner, J., Turner, R., Friston, K.J. (2001). Modeling geometric deformations in EPI time series. *NeuroImage*, *13*, 903–19.

Banse, R., Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*, 614–36.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B.V. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*, 309–12.

Boul, K.E., Funk, W.C., Darst, C.R., Cannatella, D.C., Ryan, M.J. (2007). Sexual selection drives speciation in an Amazonian frog. *Proceedings of the Royal Society B*, *274*, 399–406.

Bradley, M.M., Lang, P.J. (1994). Measuring emotion: the self-assessment Manikin and the Semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49–50.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C. (1994). Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, *18*, 192–205.

Costa, M., Braun, C., Birbaumer, N. (2003). Gender differences in response to pictures of nudes: a magnetoencephalographic study. *Biological Psychiatry*, *63*, 129–47.

Ellsworth, P.C., Scherer, K.R. (2003). Appraisal processes in emotion. In: Davidson, R.J, Scherer, K.R., Goldsmith, H.H., editors. *Handbook of Affective Sciences*. Oxford: Oxford University Press, pp. 572–95.

Ethofer, T., Anders, S., Wiethoff, S., et al. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *NeuroReport*, *17*, 249–53.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J. (1994). Statistical parametric maps in neuroimaging: a general linear approach. *Human Brain Mapping*, *2*, 189–210.

Grandjean, D., Sander, D., Pourtois, G., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*, 145–46.

Hauser, M.D. (1993). Rhesus monkey copulation calls: honest signals for female choice? *Proceedings of the Royal Society of London Series B*, *254*, 93–6.

Herbert, C., Kissler, J., Junghofer, M., Peyk, P., Rockstroh, B. (2006). Processing of emotional adjectives: evidence from startle EMG and ERPs. *Psychophysiology*, *43*, 197–206.

Jaencke, L., Wuestenberg, T., Scheich, H., Heinze, H.J. (2002). Phonetic perception and the temporal cortex. *NeuroImage*, *15*, 733–46.

Karama, S., Lecours, A.R., Leroux, J.-M., Beaudoin, G., Joubert, S., Beauregard, M. (2003). Areas of brain activation in males and females during viewing of erotic film excerpts. *Human Brain Mapping*, *16*, 1–13.

Kiflawi, M., Gray, D.A. (2000). Size dependent response to conspecific mating calls by male crickets. *Proceedings of the Royal Society of London Series B*, *267*, 2157–61.

von Kriegstein, K.V., Giraud, A.L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, *22*, 948–55.

von Kriegstein, K.V., Kleinschmidt, A., Sterzer, P., Giraud, A.L. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience*, *17*, 367–76.

Sabatinelli, D., Flaisch, T., Bradley, M.M., Fitzsimmons, J.R., Lang, P.J. (2004). Affective picture perception: gender differences in visual cortex. *NeuroReport*, *15*, 1109–12.

Smith, C.A., Lazarus, R.S. (1990). Emotion and adaptation. In: Pervin, L.A., editor. *Handbook of Affect and Social Cognition*. Mahwah, NJ: Lawrence Erlbaum, pp. 75–92.

Surguladze, A., Brammer, M.J., Young, A.W., et al. (2003). A preferential increase in the extrastriate response to signals of danger. *Neuroimage*, *19*, 1317–28.

Warren, J.D., Scott, S.K., Price, C.J., Griffiths, T.D. (2006). Human brain mechanisms for the early analysis of voices. *NeuroImage*, *31*, 1389–97.