Aus der Universitätsklinik für Psychiatrie und Psychotherapie Tübingen
Abteilung Allgemeine Psychiatrie und Psychotherapie mit Poliklinik
Ärztlicher Direktor: Prof. Dr. A. J. Fallgatter

Cerebral processing of emotional prosody — influence of acoustic parameters, arousal and the role of cross-gender interactions

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von
Sarah Wiethoff
aus Nordhorn
2010

Dekan: Professor Dr. I. B. Autenrieth

1. Berichterstatter: Professor Dr. D. Wildgruber

2. Berichterstatter: Privatdozent Dr. I. Hertrich

Für meine Familie und meine Freunde

Abbreviations

ANOVA	analysis of variance
BA	Brodman-Area
BOLD	Blood-oxygen-level-dependency
EPI	echo planar imaging
fMRI	functional Magnetic Resonance Imaging
FOV	field of view
Hz	hertz
MPRAGE	magnetization prepared rapid acquisition gradient echo
ms	milliseconds
Px	pixel
SAM	Self-Assessment-Manikin
SPM	Statistical Parametric Mapping
STG	Superior Temporal Gyrus
T	tesla
T1	longitudinal relaxation time
T2	transverse relaxation time
TE	echo-time
TR	repetition-time

Table of contents

1 General introduction

Nonverbal signals play an important role in the way humans communicate with each other. Body movements like gestures and facial expressions are only one part of it – another important factor is prosody, in the clinical context firstly defined by Monrad-Kohn (1947) as that special facility of language which creates independently from semantics different meanings via modulation of speech-rhythm, loudness, frequency and stress patterns.

Approximately, only seven percent of the information about the emotional state of a speaker are inferred from semantics, meaning the content of his words or “what” he or she says. 55 percent is conveyed by body language and the rest, impressive 38 percent, is transported via prosody, e. g. “how” one says, what he says (Mehrabian, 1972).

Therefore, prosody – and its adequate interpretation – represents a vital tool within human every-day-life.

More specifically, one can further distinguish between linguistic – stressing the semantic message or syntactic meaning (i.e. question versus statement) - and emotional prosody. Emotional prosody (sometimes also referred to as affective prosody) carries information about the intentions, the emotional state and the personality of the speaker and will be in the focus of this thesis.

But what does emotional prosody exactly consist of? So far, a lot of research has been carried out to further disentangle the contribution of different acoustic parameters to the expression of emotional prosody. Numerous scientists tried to clarify the influence and importance of single acoustic features within the creation of different emotional intonations (like for example anger, happiness, disgust, sadness and fear). One can find a summary of the partly inconsistent findings in Pittam and Scherer (1993), but to make the point one can retain an important role for the following acoustic cues in conveying emotional prosody:

a) fundamental frequency, b) intensity (also perceived as loudness) and c) temporal information like rhythm and pauses.

The following study wanted to create a link between these mainly neuropsychological studies pointing out the importance of different acoustic features in perception and expression of emotional prosody and brain imaging data detecting the neural correlates of processing emotional prosody.

Therefore, a neuroimaging experiment was designed which again targeted the voice-processing areas that were previously described as important for elaboration of emotional prosody (Grandjean et al., 2005, Ethofer et al., 2006a), but allowed at the same time to investigate the impact of basic acoustic parameters in the prosodic signal and other possible influences on these areas.

To this aim, a careful selection of the stimulus set was of high importance. Several pre-studies were conducted in order to ensure that the intended emotion within the prosodic signal was reliably recognized (categorization rating) and to assess the valence and arousal of the different stimuli (see Fig. 1). To obtain an experimental design that is balanced for the valence of the emotional information expressed by prosody, two negative (i.e. anger and fear) and two positive emotions (i.e. happiness and eroticism) were chosen in addition to neutral prosody, serving as control. This design enabled us to investigate a broad spectrum of emotional categories expressed by prosody and their impact on voice-sensitive areas.

In the first part of this thesis, differential contributions of certain acoustic parameters on increased responsiveness in right mid STG to emotional versus neutral prosody are investigated (Wiethoff et al., 2008). In this approach separate simple regression analyses of different acoustic parameters were used to partial out their effect on brain activation in voice processing regions. In addition, a multiple regression analysis with all five investigated acoustic parameters was conducted to evaluate the conjoint effect of these parameters on the hemodynamic responses in our target brain areas.

The main focus of the second part of the thesis was the investigation of effects that cannot be solely explained by differences in acoustic parameters, but depend on the behavioural relevance of the expressed emotional information (Ethofer et al., 2007). For this aim, the inclusion of stimuli spoken in an erotic tone of voice was of particular importance, since in heterosexual individuals seductively spoken speech-signals gain a high behavioural relevance (e.g. reproduction) as well as an extended meaning when different sexes talk and listen to each other. According to our hypothesis, this fact should enable us to detect behavioural as well as neurophysiological reactions that cannot be exclusively explained on the basis of different acoustic cues, as the physically identical stimuli (with the same variability of the acoustic parameters in it) do or do not gain additional information depending on the gender of speaker and listener.

In order to investigate these two aspects the following study was conducted, which is in more detail described in the two publications that form the main part of the manuscript.

A general summary that concludes the important findings of both papers will complete this doctoral thesis and can be additionally found in a German version. Since the quotations of the two publications were redundant in some cases, one common list of references is given at the end of the manuscript.

The research project was published in two articles in scientific journals (in slightly modified versions):

1) Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T. Cerebral processing of emotional prosody – influence of acoustic parameters and arousal, NeuroImage, 39: 885-893, 2008.

2) Ethofer T, Wiethoff S, Anders S, Kreifelts B, Grodd W, Wildgruber D. The voices of seduction: cross-gender effects in processing of erotic prosody. Social Cognitive and Affective Neuroscience (SCAN), 2: 2334-2337, 2007.

2 Cerebral processing of emotional prosody—influence of acoustic parameters and arousal

Sarah Wiethoff,[a,b] Dirk Wildgruber,[a,b] Benjamin Kreifelts,[a,b] Hubertus Becker,[c] Cornelia Herbert,[d] Wolfgang Grodd,[b] and Thomas Ethofer, [a,b]

[a]Department of Psychiatry, University of Tuebingen, Tuebingen, Germany
[b]Section on Experimental MR of the CNS, Department of Neuroradiology, University of Tuebingen, Tuebingen, Germany
[c]Department of Neurology, University of Tuebingen, Tuebingen, Germany
[d]Department of Psychology, University of Konstanz, Konstanz, Germany

2.1 Abstract

The human brain has a preference for processing of emotionally salient stimuli. In the auditory modality, emotional prosody can induce such involuntary biasing of processing resources. To investigate the neural correlates underlying automatic processing of emotional information in the voice, words spoken in neutral, happy, erotic, angry, and fearful prosody re presented in a passive-listening functional magnetic resonance imaging (fMRI) experiment. Hemodynamic responses in right mid superior temporal gyrus (STG) were significantly stronger for all emotional than for neutral intonations. To disentangle the contribution of basic acoustic features and emotional arousal to this activation, the relation between event-related responses and these parameters was evaluated by means of regression analyses. A significant linear dependency between hemodynamic responses of right mid STG and mean intensity, mean fundamental frequency, variability of fundamental frequency, duration, and arousal of the stimuli was observed. While none of the acoustic parameters alone explained the stronger responses of right mid STG to emotional relative to neutral prosody, this stronger responsiveness was abolished both by correcting for arousal or the conjoint effect of the acoustic parameters. In

conclusion, our results demonstrate that right mid STG is sensitive to various emotions conveyed by prosody, an effect which is driven by a combination of acoustic features that express the emotional arousal in the speaker's voice.

2.2 Introduction

In a multifaceted environment, our sensory systems are confronted with an abundance of information which contrasts with the limited processing capacity of the human brain. To cope with these limitations, stimuli of potential behavioral relevance have to be separated automatically from irrelevant ones to enable filtering of vital information. This separation can either occur by voluntary attention to certain stimulus features or in an involuntary stimulusdriven manner (Desimone and Duncan, 1995). Such automatic processing has been observed in response to novel stimuli (Näätänen, 1990) or emotionally salient stimuli (Vuilleumier and Schwartz, 2001; Vuilleumier, 2005). In the auditory domain, emotionally salient information can be expressed via modulation of speech melody (emotional prosody). Findings obtained from auditory evoked potentials indicate differential processing of emotionally and neutrally spoken vowels around 400 ms after stimulus onset (Alter et al., 2003). In previous functional magnetic resonance imaging (fMRI) experiments, it has been demonstrated (Grandjean et al., 2005; Ethofer et al., 2006a; Beaucousin et al., 2007) that voice-sensitive regions (Belin et al., 2000) in the associative auditory cortex in the middle part of the superior temporal gyrus (mid STG) adjacent to superior temporal sulcus respond stronger to happy and angry than to neutral intonations.

This response pattern was found independently of whether the subjects were instructed to attend to emotional prosody (Ethofer et al., 2006a; Beaucousin et al., 2007) or some other feature of the presented stimuli, such as emotional semantics (Ethofer et al., 2006a) or speakers' gender (Grandjean et al., 2005). These findings suggest that activity in mid STG regions is mainly driven by bottom-up mechanisms which rely on features of the presented stimuli and is relatively unaffected by top-down processes that focus the attention of the subjects on certain stimulus characteristics.

So far, it is unclear which stimulus-bound features mediate these bottom-up mechanisms and whether stronger responses to emotional as compared to neutral prosody can be explained by differences in one single acoustic parameter. To minimize such biases, previous studies investigating the neural correlates underlying perception of emotional prosody employed stimuli which were matched for basic parameters, such as acoustic energy (Grandjean et al., 2005) or maximum peak intensity (Ethofer et al., 2006a).

Obviously, such approaches are limited by the fact that emotional information in the voice is transmitted via certain acoustic features (Banse and Scherer, 1996) and matching emotional stimuli for all possible acoustic parameters to neutral stimuli would presumably remove the emotional information conveyed by prosody.

Therefore, an alternative approach was employed in the present study:

First, the impact of single acoustic parameters or the conjoint effect of a set of acoustic parameters was evaluated in simple or multiple regression analyses, respectively. After removing all the variance correlating with the parameters in question, we tested whether the respective regression residuals still show the effect of stronger responsiveness to emotional than neutral prosody.

Here, we used event-related fMRI to investigate the neuronal correlates underlying automatic processing of a broad variety of prosodic categories expressing highly arousing emotional information. To this end, we presented words spoken in neutral, happy, erotic, angry, or fearful intonation in a passive listening paradigm. Based on results of previous neuroimaging experiments (Grandjean et al., 2005; Ethofer et al., 2006a; Beaucousin et al., 2007), we hypothesized that voice-sensitive regions in mid STG show stronger blood oxygen level dependent (BOLD) responses to emotional intonations expressing high emotional arousal than to low arousing neutral intonations. To evaluate whether stronger responsiveness of this region can be explained by the perceived emotional arousal or by basic acoustic parameters, such as mean intensity, variability (standard deviation) of intensity, mean fundamental frequency, variability of fundamental frequency, or stimulus duration, separate simple regression analyses were carried out and the resulting regression residuals of emotional and neutral trials were statistically compared. Furthermore, a multiple regression analysis, including all five acoustic parameters investigated here, was conducted and regression residuals of emotional and neutral trials obtained in this analysis were compared to evaluate the conjoint effect of these parameters on event-related responses of voice-sensitive regions in mid STG.

2.3 Material and methods

2.3.1 Subjects

Twenty-four right-handed, healthy, native, or highly proficient German speaking participants (12 female, 12 male, mean age 25.1 years) were included in an fMRI study. None of the subjects previously participated in an fMRI experiment on emotional processing. Right-handedness was assessed using the Edinburgh Inventory (Oldfield, 1971). The study was approved by the Ethical Committee of the University of Tuebingen and conducted according to the Declaration of Helsinki.

2.3.2 Stimuli

126 adjectives and nouns that had been judged by 45 native speakers of German according to their emotional word content on a 9-point self-assessment manikin scale (SAM, Bradley and Lang, 1994) along the dimensions of valence (ranging from 1=highly positive to 9=highly negative) and arousal (ranging from 1=very calming to 9=highly arousing) were selected for the present study. These words were rated as emotionally neutral (mean valence values between 4 and 6) and low-arousing (mean arousal ratings < 4). Six professional actors (3 female, 3 male) pronounced these words in either neutral, happy, erotic, angry, fearful, sad, or disgusting intonation (18 words each). All stimuli were normalized to the same maximum peak intensity. To ensure that stimuli represent the emotional category

intended by the actors and that emotional intonations are balanced for arousal, two prestudies were conducted. In the first prestudy, 10 healthy volunteers (mean age 24 years, 4 female, 6 male) were instructed to assign the emotional prosody to one of seven possible emotional intonations (neutral,

happy, sexually alluring, angry, fearful, sad, or disgusting). Stimuli that were assigned to the intended emotional category by at least 70% of the volunteers were additionally evaluated in a second prestudy. Since we aimed at presenting an equal number of stimuli with positive and negative valence in the fMRI study, stimuli of only two negative emotion categories were included in this prestudy. We selected fearful and angry intonations because more stimuli of these categories were correctly identified by more than 70% of the participants than for the categories of sadness and disgust. In the second prestudy, 20 healthy volunteers (mean age 25.2 years, 10 female, 10 male) judged the arousal of the expressed emotional prosody on a 9-point SAM scale (see Fig. 1).

The final stimulus set employed in the fMRI experiment comprised 25 nouns and 25 adjectives and was balanced for arousal of the emotional intonations (happy: 5.65 ± 0.64, erotic: 5.89 ± 0.36, angry: 5.72 ± 0.47, fearful: 5.60 ± 1.09, neutral: 2.45 ± 0.45; mean values ± standard deviation, see Fig. 1 and Table 1), word frequency in spoken language (neutral: 0.83 ± 0.93, happy: 0.7 ± 0.55, erotic: 0.87 ± 0.76, angry: 0.86 ± 0.74, fearful: 0.86 ± 0.74; mean logarithmic word frequency ± standard deviation as assessed by the Mannheim Corpus of the Institut für Deutsche Sprache, Mannheim), number of syllables, and gender of the speaker.

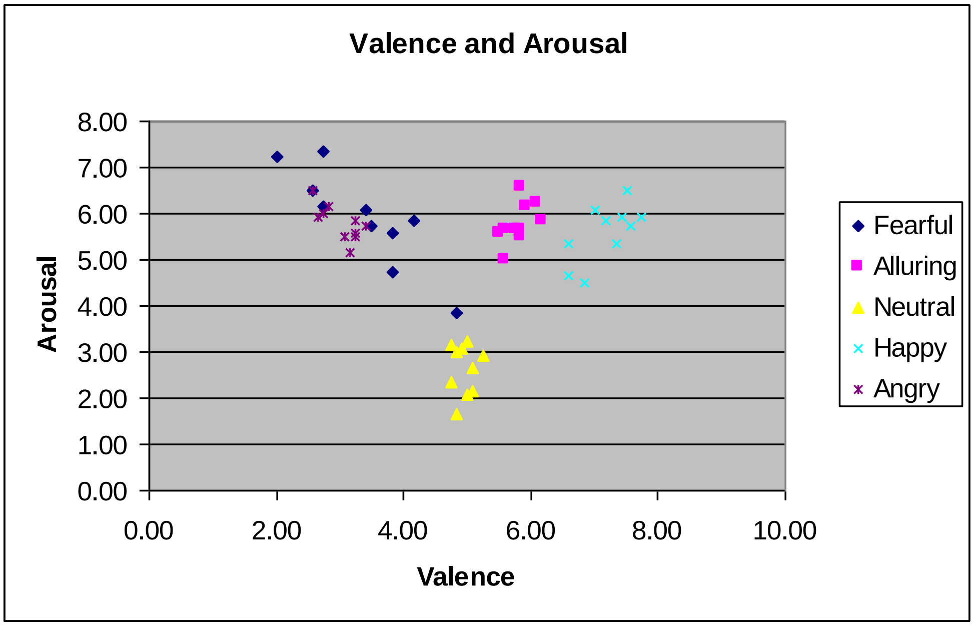

Fig. 1: Mean arousal- and valence-ratings (rated by 20 healthy individuals on a nine-point-self-assessment-manikin-scale) of all selected stimuli sorted by its emotional category.

Table 1: Basic acoustic parameters of the stimulus material (Mean ± standard deviation)

	Mean I (dB)	Variability of I * (%)	Mean FO (Hz)	Variability of FO * (%)	Duration (s)
Angry	75 ± 3	22 ± 4	174 ± 27	16 ± 9	0.78 ± 0.27
Fearful	74 ± 4	20 ± 3	185 ± 16	12 ± 5	0.79 ± 0.24
Neutral	75 ± 2	19 ± 1	156 ± 19	10 ± 4	0.71 ± 0.20
Happy	77 ± 3	22 ± 2	182 ± 20	23 ± 7	0.71 ± 0.17
Erotic	75 ± 1	16 ± 1	138 ± 13	16 ± 10	1.24 ± 0.33

dB = decibel, Hz = Hertz, s = seconds

*Variability (SD) of intensity (I) and fundamental frequency (F0) were normalized to the mean of the respective stimuli

2.3.3 Analysis of acoustic parameters

For each stimulus, the mean intensity (I) and variability of intensity (standard deviation (SD) of I, normalized to mean I), mean fundamental frequency (F0), and variability of F0 (SD of F0, normalized to the mean F0) were calculated from I and F0 contours determined for the whole stimulus duration using Praat software (version: 4.3.20, http://www.praat.org, Boersma, 2001). Mean I, SD of I, mean F0, SD of F0, and the duration of each stimulus (see Table 1) were used in separate regression analyses (see below) to test whether hemodynamic responses to stimuli of different emotional categories are still significantly stronger than responses to neutral stimuli after eliminating the variance correlating with these parameters.

2.3.4 Experimental design

The fMRI experiment consisted of two sessions. In both sessions, the same 50 stimuli were presented binaurally via magnetic resonance-compatible headphones with piezoelectric signal transmission (Jäncke et al., 2002). 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 to 12 s. To minimize explicit processing of emotional prosody, the participants were told that the experiment was conducted to examine differences in the perception of adjectives and nouns. Subjects were instructed that their only task was to listen to the presented stimuli.

2.3.5 Image acquisition

Structural and functional imaging data were acquired using a 3 T-whole body scanner (Siemens TRIO, Erlangen, Germany). A magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence was employed to acquire high-resolution (1×1×1 mm^3) T1-weighted structural images (TR=2300 ms, TE=3.93 ms, TI=1100 ms). Functional images were obtained using a multislice echo planar imaging (EPI) sequence (25 axial slices acquired in descending order, slice thickness 4 mm + 1 mm gap, TR=1.83 s, TE=40 ms, field of view (FOV)=192×192 mm^2, 64×64 matrix, flip angle=90°, bandwidth 1906 Hz/Px). For offline correction of EPI image distortions, a static fieldmap (36 slices acquired in descending order, slice thickness=3 mm + 1 mm gap, TR=400 ms, TE (1)=5.19 ms, TE (2)=7.65 ms, FOV=192×192 mm^2, 64×64 matrix) was acquired prior to the functional measurements.

2.3.6 Image analysis

The first five fMRI volumes were discarded from further analysis to exclude measurements that preceded T1 equilibrium. Functional images were analyzed using statistical parametric mapping software (SPM2, Wellcome Department of Imaging Neuroscience, London, UK). Preprocessing steps comprised realignment to the first volume of the time series, unwarping by use of a static field map (Andersson et al., 2001), normalization into MNI space (Montreal Neurological Institute, 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) in which separate regressors were defined for each trial using a stick function convolved with the hemodynamic response function. Events were time-locked to stimulus onset. To remove low frequency components, a high-pass filter with a cutoff-frequency of 1/256 Hz was used. Serial autocorrelations were accounted for by modeling the error term as a first-order autoregressive process with a coefficient of 0.2 (Friston et al., 2002) plus a white noise component (Purdon and Weisskoff, 1998). A subtraction analysis was employed to evaluate which brain regions respond stronger to emotional than to neutral intonations.

To examine whether hemodynamic responses in the right mid STG are subject to repetition suppression effects, parameter estimates of the most significantly activated voxel in this region as defined by the contrast (emotional > neutral prosody) were submitted to a two-factorial analysis of variance (ANOVA) with emotion (emotional, neutral) and repetition (first fMRI session, second fMRI session) as within-subject factors. Statistical inference was based on second-level random effects analyses. Activations are reported descriptively at a height threshold of $p<0.001$ and an extent threshold of $k>12$ voxels. 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.

Where available, probabilistic cytoarchitectonic maps were used to label functional activations within the auditory cortex using the SPM anatomy toolbox (Eickhoff et al., 2005). Otherwise, functional activations were assigned to macroscopic structures as determined by automatic anatomic labeling (Tzourio-Mazoyer et al., 2002).

Separate simple regression analyses were conducted to investigate in which brain regions event-related responses covary with mean I, variability of I, mean F0, variability of F0, and duration and/or arousal of the stimuli. Furthermore, event-related responses from our region of interest in right mid STG (as defined by the subtraction analysis emotional > neutral prosody) were extracted to evaluate whether enhanced responsiveness to emotional prosody in this brain region can be solely explained on the basis of one single acoustic parameter or emotional arousal. Emotional arousal of each stimulus was determined in a prestudy (see above).

To examine the impact of the acoustic parameters and emotional arousal on hemodynamic responses of our region of interest, correlation coefficients obtained for each subject were Fisher Z transformed and submitted to a random effects analysis based on a one-sample t-test. Mean correlation coefficients for each acoustic parameter and emotional arousal were calculated by backtransformation of mean Fisher Z scores. Since the duration of the stimuli is both dependent on the number of syllables they are composed of and the rate of articulation at which these syllables are spoken, an additional analysis was carried out to evaluate whether the number of syllables had a significant impact on the relation between mid STG responses and stimulus duration. To this end, a simple regression analysis with number of syllables as independent variable and event-related responses in right mid STG as dependent variable was performed. Regression residuals obtained from this analysis were subsequently correlated with mean syllable duration (i.e. stimulus duration divided by number of syllables) and resulting Fisher Z scores were compared by paired t-tests to Fisher Z scores obtained from correlation analyses without correction for possible influences of the number of syllables.

The regression residuals obtained from the simple regression analyses were subsequently tested in separate paired t-tests to investigate whether responses to emotional prosody are still significantly stronger than responses to neutral prosody after neutralization of possible contributions of single acoustic parameters or emotional arousal. To evaluate whether stronger responses to emotional as compared to neutral prosody can be explained by the conjoint effect of the acoustic parameters, a multiple regression analysis including all five acoustic parameters investigated here was carried out. Accordingly, the regression residuals were submitted to paired t-tests.

2.4 Results

Comparison of hemodynamic responses to emotional and neutral prosody revealed a cluster with two distinct maxima in the right temporal lobe, which were located within Heschl's gyrus (MNI coordinates: x=48; y=−24; z=3; Z score=4.06, cytoarchitectonic subdivision TE 1.1 as defined by probabilistic maps obtained from Morosan et al., 2001 and Rademacher et al., 2001) and in the mid STG (MNI coordinates: x=63; y=−12; z=0; Z score=3.72, k=99, $p<0.05$ corrected, see Fig. 2a–c).

Other brain regions showing stronger activations to emotional than to neutral intonations included the left temporal pole (MNI coordinates: −36 6 −18, Z score=4.10, k=48), and hypothalamus (MNI coordinates: x=9; y=0; z=−6; Z score=3.74, k=44). Post hoc paired t-tests on the contrast estimates obtained separately for each of the five prosodic categories revealed that responses of right mid STG to all emotional intonations were significantly stronger than responses to neutral intonations (angry vs. neutral $t(23)=2.63$, fearful vs. neutral $t(23)= 2.71$, happy vs. neutral $t(23)=4.41$, erotic vs. neutral $t(23)=4.86$, all $p<0.01$, one-tailed, see Fig. 2d).

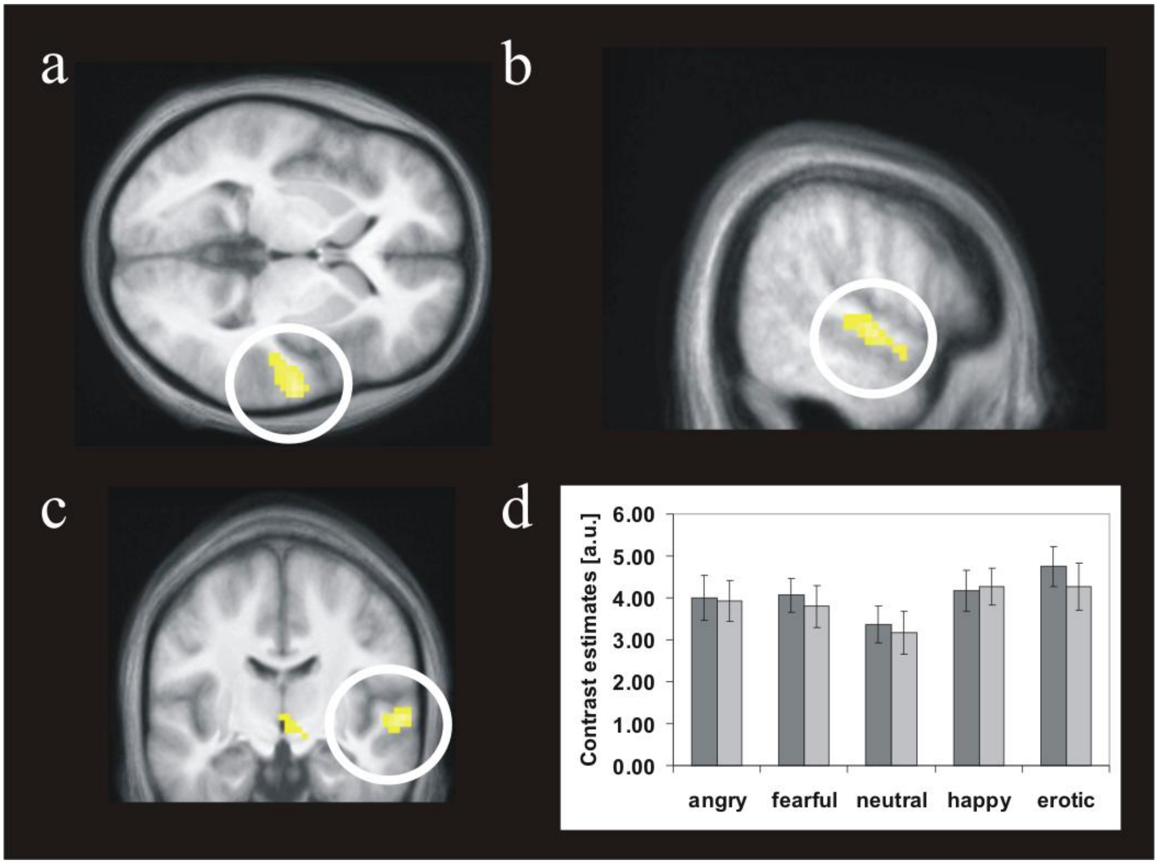

Fig. 2: Stronger activation to emotional than to neutral prosody in right mid STG shown on transversal (a), saggital (b) and coronal (c) slices, thresholded at $p<0.05$ (corrected). Separate contrast estimates (d) for the five prosodic categories obtained from right mid STG (MNI-coordinates: x = 63; y = -12; z = 0). Dark and light grey bars represent separate parameter estimates of the first and the second fMRI session, respectively.

The two-factorial ANOVA with emotion and repetition as within-subject factors and brain responses in the right mid STG as dependent variable revealed a significant main effect of emotion ($F(1,23)=19.44$, $p<0.001$), but no main effect of repetition ($F(1,23)= 0.63$, $p=0.44$). Moreover, no interaction between emotion and repetition was found ($F(1,23)= 0.01$, $p=0.94$). Parameter estimates obtained for the two imaging runs were similar for all presented prosodic categories with a tendency for slightly diminished responses during the second run (see Fig. 2d). Post hoc paired t-tests did not reveal any significant differences between hemodynamic responses obtained for the first as compared to the second run (angry: $t(23)=0.19$, fearful: $t(23)=0.91$, neutral: $t(23)=0.53$, happy: $t(23)=-0.26$, erotic: $t(23)=1.51$, all $p>0.05$, one-tailed).

Brain regions showing a significant positive or negative linear relationship between hemodynamic responses and mean I, variability of I, mean F0, variability of F0, stimulus duration, and emotional arousal are rendered onto a standard brain template in Figs. 3a–f, respectively.

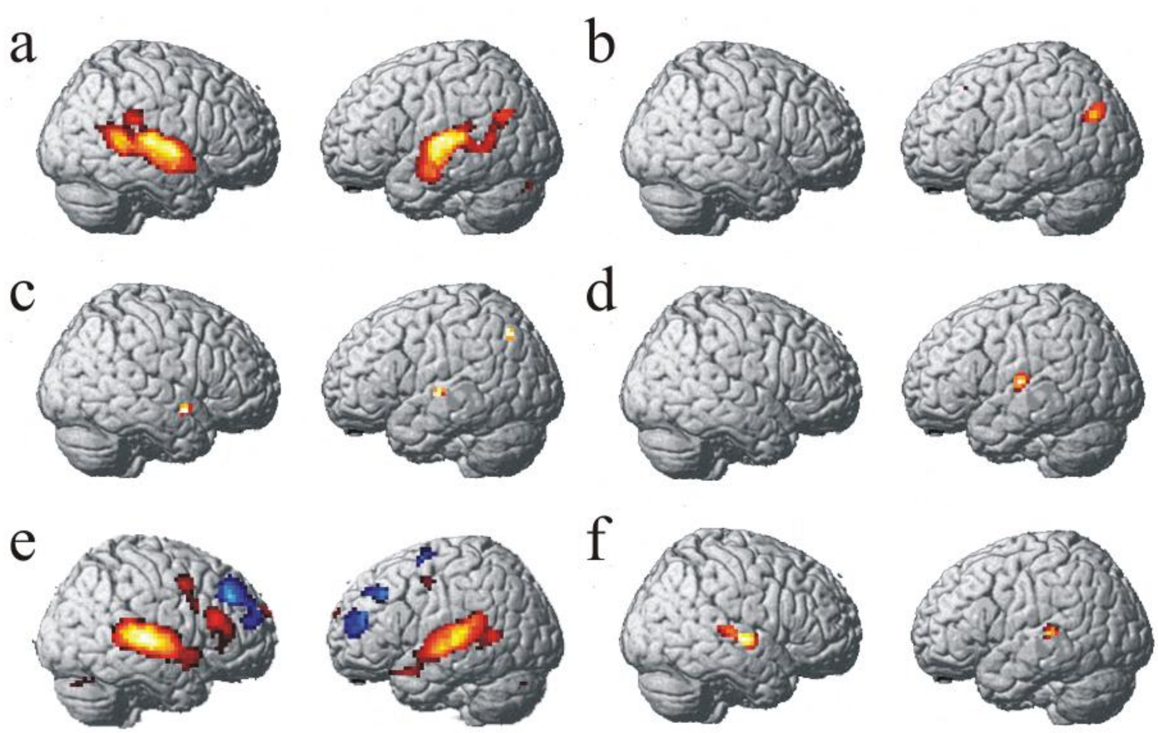

Fig. 3: Brain regions showing positive (yellow/red clusters) or negative (light/dark blue clusters) correlations of their hemodynamic responses with (a) mean intensity, (b) variability of intensity, (c) mean fundamental frequency, (d) variability of fundamental frequency, (e) duration, and (f) arousal of the stimuli.

Hemodynamic responses in large parts of both temporal lobes exhibited a positive correlation with mean I and stimulus duration. Two small clusters in the right temporal pole and left superior temporal cortex demonstrated a positive relationship of their hemodynamic responses with mean F0. For variability of F0, a positive relationship with hemodynamic responses in left Heschl's gyrus was found. Emotional arousal showed a linear relationship with hemodynamic responses of bilateral mid STG. Other brain regions showing either a positive or negative linear relationship between their hemodynamic responses with one of the acoustic parameters investigated here are given in Table 2.

Table 2: Correlations between acoustic parameters and event-related hemodynamic responses

Anatomical definition	MNI coordinates	Z score	Cluster size
Mean intensity: positive correlations			
Right superior/middle temporal gyrus	60 -3 -3	5.76	1072
Left superior/middle temporal gyrus	-63 -15 -3	5.67	820
Cerebellum, vermis	-3 -72 -21	3.72	99
Left angular gyrus	-42 -63 27	3.59	82
Left posterior cingulum	-6 -39 30	3.49	37
Left hippocampus	-18 -30 -9	3.86	31
Variability of intensity: positive correlations			
Bilateral precuneus	-12 -60 -18	3.70	233
Left angular gyrus	-42 -72 -27	3.88	81
Left middle frontal gyrus	-21 18 45	3.85	43
Left parahippocampal gyrus	-27 -42 9	3.78	37
Mean pitch: positive correlations			
Left angular gyrus	−42 −66 48	3.44	15
Left superior temporal gyrus	−51 −15 0	3.29	15
Right temporal pole	60 3 −12	3.73	13
Variability of pitch: positive correlations			
Left Heschl's gyrus	−51 −15 9	3.91	82
Duration: positive correlations			
Left superior/middle temporal gyrus	−51 −27 3	6.63	1139
Right superior/middle temporal Gyrus	45 −30 9	7.39	1087
Right inferior frontal gyrus	57 27 18	4.84	142
Right cerebellum	27 −75 −30	3.99	83
Right precentral gyrus	48 6 45	4.34	74

Left cerebellum	−24 −78 30	4.23	45
Right superior frontal gyrus, medial Part	3 66 27	3.62	25
Left precentral gyrus	−48 −3 −48	3.46	19
Duration: negative correlations			
Right middle frontal gyrus	42 39 36	4.56	222
Left middle frontal gyrus	−45 48 18	4.17	118
Left middle frontal gyrus	−45 30 39	3.71	59
Left superior frontal gyrus	−18 −3 69	3.58	36
Arousal: positive correlations			
Right mid STG	66 −12 0	4.12	145
Left mid STG	−54 −33 6	3.43	28
Hypothalamus	0 −9 3	3.47	17

In Figs. 4a–f, event-related responses extracted from our region of interest in right mid STG and averaged over subjects are plotted against mean I, variability of I, mean F0, variability of F0, stimulus duration, and emotional arousal. Activity in right mid STG showed a considerable correlation with mean I (mean $r=0.19$, $t(23)=7.0$, $p<0.001$, two-tailed) and duration (mean $r=0.32$, $t(23)=8.8$, $p<0.001$, two-tailed). The correlation between mid STG responses and duration was significantly smaller after correction for possible influences of the number of syllables (paired $t(23)=7.8$, $p<0.001$, one-tailed), but still significantly different from zero across subjects (mean $r=0.16$, $t(23)=6.1$, $p<0.001$, two-tailed) indicating that the linear relationship between BOLD responses and duration was attributable to both varying number of syllables and rate of articulation. Hemodynamic responses in right mid STG were only weakly correlated with mean F0 (mean $r=0.10$, $t(23)=2.8$, $p<0.05$, two-tailed) and variability of F0 (mean $r=0.09$, $t(23)=4.1$, $p<0.001$, one-tailed). No significant correlation was found for variability of I (mean $r=-0.05$, $t(23)=-1.8$, $p=0.08$, two-tailed). Activity in right mid STG showed a significant correlation with emotional arousal (mean $r=0.14$, $t(23)=5.6$, $p<0.001$, two-tailed).

Regression residuals obtained from simple regression analyses between parameter estimates obtained from the right mid STG and acoustic parameters were significantly bigger for trials with emotional prosody than for trials with neutral prosody (for all five acoustic parameters, $t(23)>2.0$, $p<0.05$, one-tailed). However, comparison of emotional and neutral trials did not yield any statistical differences for regression residuals obtained from the simple regression analysis between parameter estimates of right mid STG and emotional arousal ($t(23)=0.9$, $p=0.19$, one-tailed). Similarly, the multiple regression analysis including all five acoustic parameters yielded regression residuals which were not significantly different for emotional and neutral prosody ($t(23)=0.7$, $p=0.25$, one-tailed).

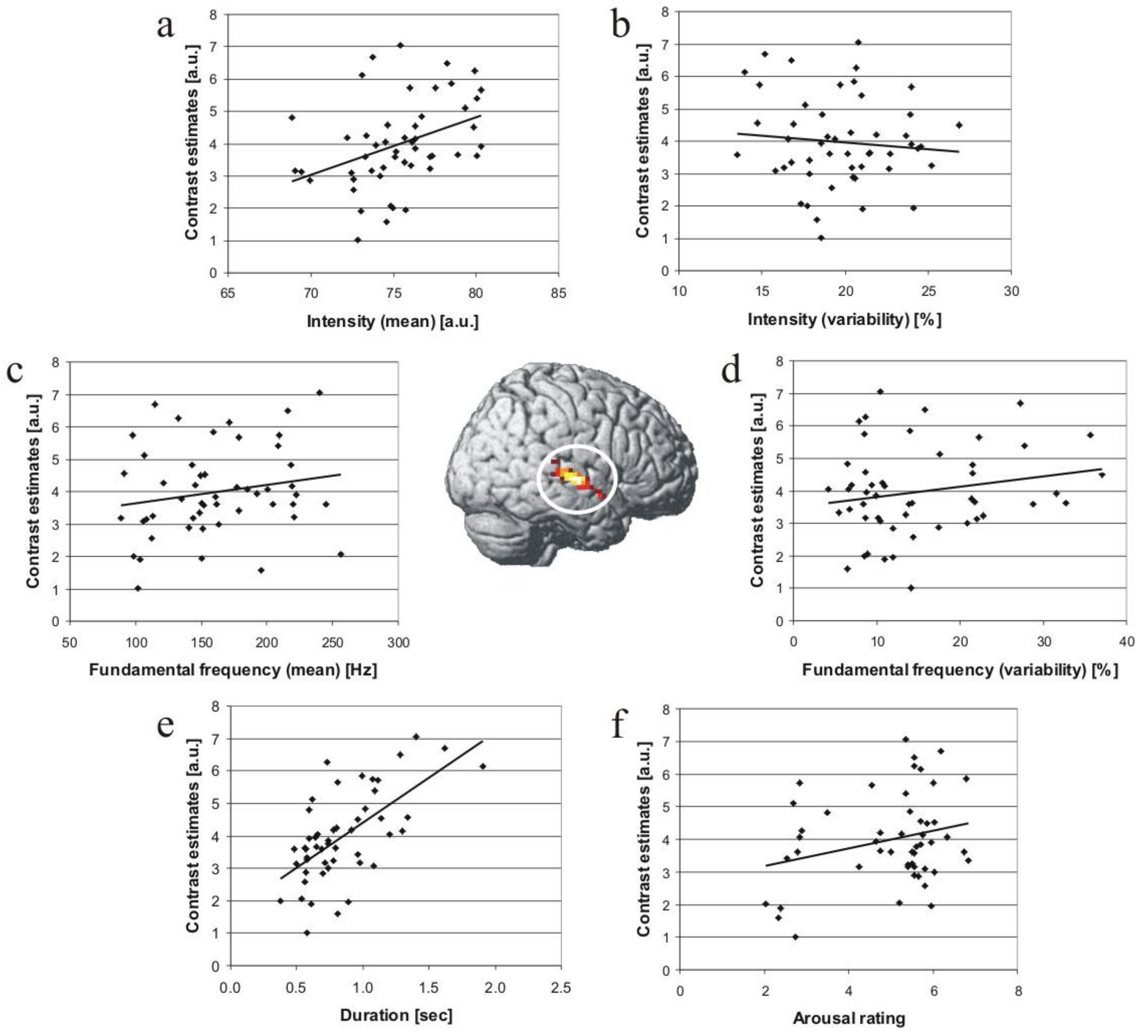

Fig. 4: Mean responses in right mid STG to stimuli of all five prosodic categories are plotted vs. (a) mean intensity, (b) variability of intensity, (c) mean fundamental frequency, (d) variability of fundamental frequency, (e) duration, and (f) arousal of the stimuli. Regression lines represent average regression slopes and y-intercepts obtained from individual subjects (a.u. = arbitrary units).

2.5 Discussion

The present study was conducted to investigate the neural correlates underlying automatic processing of emotional prosody. In particular, we wanted to test whether our region of interest in right mid STG shows increased responses to a broad variety of emotional intonations. Furthermore, we wanted to clarify whether stronger responses within this brain region can be explained on the basis of emotional arousal, a single acoustic parameter, or the conjoint effect of mean intensity, variability of intensity, mean fundamental frequency, variability of fundamental frequency, and stimulus duration.

2.5.1 Comparison of brain responses to emotional and neutral prosody

In agreement with our hypothesis, all emotional intonations induced stronger hemodynamic responses than neutral intonations within the cortex of the right mid STG. Thus, our results confirm previous reports on that region showing increased responsiveness to prosodically expressed happiness and anger (Grandjean et al., 2005; Ethofer et al., 2006a; Beaucousin et al., 2007) and extend these findings to a broad spectrum of emotionally arousing prosody, including intonations of both positive and negative valence as well as intonations signalling withdrawal or approach.

Furthermore, demonstration of this effect in an fMRI paradigm without task instructions underlines its stimulus-driven nature and is in agreement with previous suggestions that enhanced responses in mid STG cortices to emotional prosody occur independent of cognitive demands (Ethofer et al., 2006a). In addition to the right mid STG, increased hemodynamic responses to emotional as compared to neutral prosody were also found in right primary auditory cortex, left temporal pole, and hypothalamus.

Enhanced responses to vocal emotional information in these regions are in keeping with frequent implication of the temporal pole in social and emotional processing (for a review see Olson et al., 2007) and mediation of autonomic responses to emotional stimuli by the hypothalamus (LeDoux, 1993). However, since our a priori predictions were restricted to the right mid STG these results can only be reported descriptively here.

No significant effect of stimulus repetition or interaction between emotion and repetition was found in right mid STG. This negative result requires careful interpretation. However, our findings suggest that possible suppression effects induced by one-time repetition are small and of comparable extent for neutral and emotional intonations. This finding is in line with previous reports obtained from the visual domain indicating that interactions between emotion and repetition occur only if the stimulus is a target in the context of a cognitive task (Ishai et al., 2004).

2.5.2 Correlation between acoustic parameters, emotional arousal and brain responses

Simple regression analyses between event-related responses and acoustic parameters yielded large clusters in bilateral temporal cortex for mean volume intensity and duration of the stimuli. These findings are in agreement with previous findings (Mathiak et al., 2002; Schall et al., 2003) demonstrating widespread activation clusters during odd-ball designs in bilateral auditory cortex to stimuli that are deviant in intensity or duration.

Inspection of responses in right mid STG revealed a fairly linear relationship within the investigated intensity range. These results converge with previous findings demonstrating a linear relationship of BOLD response and sound intensity for a large range of intensity (Jäncke et al., 1998, Langers et al., 2007). Event-related responses in right mid STG also showed a linear relationship with stimulus duration. A separate regression analysis demonstrated that this effect was partly attributable to the varying number of syllables of the stimuli. However, the regression residuals of this analysis did still show a considerable correlation with duration indicating that this effect was also driven by the rate of articulation resulting in varying durations of the individual syllables.

Event-related responses in two small clusters in bilateral anterior temporal cortex showed a linear increase with mean F0 which is in line with results from a previous study demonstrating that the anterior temporal cortex responds stronger to female voices with high F0 than to male voices with low F0 (Sokhi et al., 2005). Moreover, it has been demonstrated that the anterior temporal cortex is implicated in speaker identification (von Kriegstein and Giraud, 2004), a function which partly relies on decoding of fundamental frequency.

Thus, our results are in agreement with the suggestion that pitch processing involves a ventral "what"-processing stream for auditory object analysis

(Rauschecker and Tian, 2000; Zatorre et al., 2004; Barrett and Hall, 2006). Variability of fundamental frequency and intensity correlated with activity in left primary auditory cortex and left angular gyrus, respectively. It has been suggested that the lateral part of Heschl's gyrus serves as a "pitch center" (Griffiths, 2003). Furthermore, our results are in line with previous findings demonstrating that left auditory cortex is implicated in detection of acoustically deviant word stimuli (Tervaniemi et al., 2006). Stronger involvement of left hemispheric brain regions in decoding of these parameters might reflect a specialization of the auditory system in the left hemisphere for detecting rapid changes of the acoustic signal in the time-domain (Jäncke et al., 2002; Zatorre et al., 2002; Joanisse and Gati, 2003; Meyer et al., 2005; Reiterer et al., 2005; Mottonen et al., 2006).

Simple regression analysis between brain responses and emotional arousal ratings of the stimuli revealed two clusters in bilateral mid STG. The cluster in right mid STG was strikingly similar in location and extent to the cluster identified by contrasting hemodynamic responses of emotional and neutral prosody. This finding lends support to the notion that enhanced responsiveness of this region is driven by the emotional arousal of the speaker conveyed by prosody.

2.5.3 Effects of acoustic parameters and emotional arousal on activation of right mid STG

Event-related responses in right mid STG were significantly correlated with mean intensity, mean fundamental frequency, variability of fundamental frequency, and stimulus duration. This sensitivity to a variety of acoustic parameters is in agreement with previous suggestions that this brain region is involved in the detailed analysis of voice information (Warren et al., 2005, 2006).

To evaluate whether differences in these parameters can explain the enhanced responsiveness to emotional prosody, we tested whether this effect is still significant after removing all variance correlated with these acoustic parameters by separate simple regression analyses.

For all five acoustic parameters investigated here, regression residuals of emotional trials were significantly bigger than regression residuals of neutral intonations indicating that increased responsiveness of right mid STG to emotional prosody cannot be solely explained on the basis of one of these acoustic parameters.

However, the regression residuals obtained from the multiple regression analysis including all five acoustic parameters did not show any significant differences between emotional and neutral prosody. Similarly, the increased responsiveness to emotional prosody was abolished after correction for emotional arousal of the stimuli. It has been shown that emotional arousal covaries with a number of acoustic features (reviewed in Scherer, 2003). Thus, we suggest that stronger responses of right mid STG to emotional relative to neutral intonations are due to a combination of acoustic cues which express the emotional arousal in the speaker’s voice.

2.5.4 Limitations and consequences for future studies

Interpretation of the results of the present study should acknowledge the following limitations with respect to the employed stimuli, experimental design, data acquisition, and data analysis.

The stimuli employed in the present study represent vocal expressions of emotion portrayed by professional actors which facilitate acquisition of a highly controlled stimulus set. However, we cannot exclude that our actors overexpressed certain emotional cues and missed more subtle ones which might be only present during natural expression of emotion (Scherer, 1986). Furthermore, many acoustic parameters of natural stimuli are correlated with each other making it difficult to fully disambiguate their contribution to the measured brain responses in regression analyses. An alternative approach is to systematically manipulate the stimulus material and investigate main effects and interactions of certain acoustic features in factorial designs (e.g. Lattner et al., 2005). However, interpretation of results obtained in such factorial designs is burdened with the fact that they rely on brain responses to artificial stimulus material one never encounters in real life situations questioning to some extent their biological relevance.

We feel that regression analyses of a natural stimulus set and factorial analyses of systematically manipulated stimuli represent complementary approaches to investigate acoustic processing in future studies.

The current study aimed at investigating automatic processing of emotional prosody. Therefore, the experimental design did not include any behavioural task which might have introduced unpredictable top-down modulation of neural responses in the auditory cortex (Crinion et al., 2003). A disadvantage of such passive-listening designs is the lack of behavioural control indicating that subjects sufficiently comprehended the auditory information in the presence of scanner noise. However, part of the stimulus material used here was already employed in a previous study demonstrating that judgements of prosodic cues inside and outside the scanner are highly correlated (Ethofer et al., 2006b). Furthermore, none of the subjects reported difficulties in comprehending the presented stimuli and all subjects reported after scanning that most of the stimuli were spoken in an emotional tone of voice. Interference between scanner noise and stimulus presentation can be reduced by application of “sparse sampling” techniques in which stimuli are presented during a silent period (Hall et al., 1999). However, such imaging techniques have a low temporal resolution with only a few data points per minute making them less optimal for fMRI studies using event-related designs such as ours. In principle, it is possible to increase the temporal resolution of sparse sampling methods by repeated presentation of the same stimuli and systematic jittering of the acquisition volume relative to the stimulus onset. Unfortunately, this results in a massive increase of scanning time and induces the problem that one has to reconstruct the time course of the BOLD response from different repetitions. Possible reductions of the BOLD response due to repetition suppression, however, might render such reconstructions imprecise (Henson, 2003). Therefore, we decided for a continuous acquisition of fMRI data. However, recent methodological advancements using clustered temporal acquisition might combine silent stimulus presentation with improved temporal resolution (Schmidt et al., in press; Zaehle et al., 2007).

Finally, the regression analyses applied in the present study were restricted to linear relationships between the BOLD response and acoustic parameters. Thus, brain regions showing a strong nonlinear behavior of their responses are missed by the analyses of the present study. While for sound intensity a linear increase of the auditory BOLD response has been previously demonstrated for a large range of intensities (Langers et al., 2007), it is to be expected that other acoustic features might induce strongly non-linear responses in the auditory cortex due to possible saturation effects at the neural level (London and Hausser, 2005) or non-linear neurovascular coupling (Buxton et al., 2004). Inspection of mean responses to stimuli with durations of more than 1.2 s (see Fig. 4e) suggests that the dependency between BOLD response duration might exhibit a non-linear behavior in this range. It is obvious that BOLD responses cannot linearly increase to infinite values with increasing duration of the stimuli. However, too few data points were acquired in this range of stimulus durations to rigorously investigate whether saturation effects are already detectable for stimulus durations of about 1.2 s.

2.6 Conclusion

Prosody of all four emotional categories investigated in the present study induced stronger responses than neutral prosody in right mid STG indicating that this brain region is sensitive to a broad variety of behaviourally relevant information expressed by prosody. Demonstration of this effect in a passive-listening paradigm underlines its stimulus-driven nature and confirms previous suggestions (Ethofer et al., 2006a) that this effect occurs irrespective of task instructions.

Event-related responses in right mid STG were significantly correlated with several acoustic parameters. None of these parameters alone could explain the stronger responsiveness of this brain region to emotional prosody. However, both emotional arousal and the conjoint effect of the five acoustic parameters investigated here were sufficient to explain stronger responses in right mid STG to emotional prosody suggesting that this effect is driven by the interplay of acoustic cues which express emotional arousal in the speaker's voice.

2.7 Acknowledgments

Sarah Wiethoff was supported by a grant of the Studienstiftung des Deutschen Volkes. The study was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 550-B10) and by the Junior Science Program of the Heidelberger Academy of Sciences and Humanities.

3 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

3.1 Abstract

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.

3.2 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.

3.3 Material and Methods

3.3.1 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 piezoelectric 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 interstimulus interval ranged from 9–12 s.

3.3.2 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 highresolution (1*1*1mm^3) T1-weighted structural images (TR=2300 ms, TE=3.93 ms, TI=1100 ms).

3.3.3 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 (10mm 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.

3.4 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 5A). 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 5B). 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.

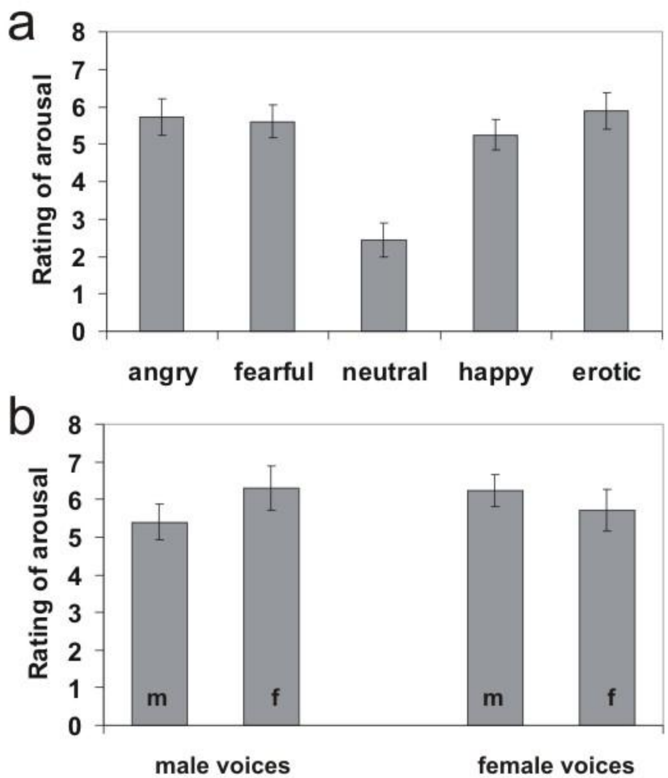

Fig.5: (A) Rating of emotional arousal of the five prosodic categories on a 9-point scale ranging from 0 (totally unexciting) to 9 (extremely exciting). (B) Rating of emotional arousal of erotic prosody dependent on the gender of speaker and listener (m=male listener, f=female listener). Error bars represent standard errors of the mean.

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 6A). 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 3).

Table 3: 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.

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 6C, angry vs neutral $T(23) > 2.63$,
fearful vs neutral $T(23)=2.71$, happy vs neutral $T(23)=4.41$, 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 6D) 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).

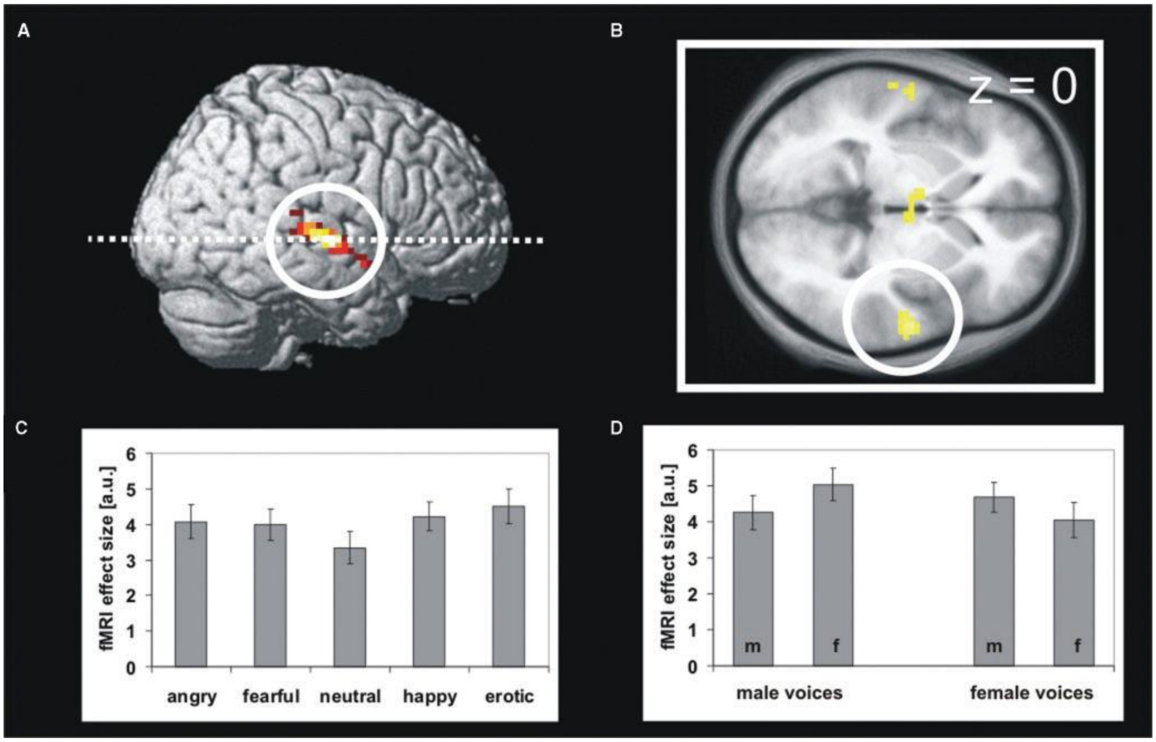

Fig. 6: (A) Brain regions showing stronger activations to emotional than neutral prosody rendered on a right hemisphere of a standard brain and (B) a transversal slice (z=0) of the mean T1 image obtained from the normalized brains of the fMRI participants. The activation cluster in the right mid STG 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.

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 where 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.

4 General summary

This fMRI-study on emotional prosody was performed in order to better understand the impact of five basic acoustic parameters on brain activation within voice processing modules of human auditory cortex. Another aim was to clarify whether basic acoustic parameters alone are sufficient to explain the increased responses in these areas in reaction to emotional compared to neutral prosody and to detect possible gender-dependent effects that are down to the behavioural relevance of the expressed emotional information.

As one of the few passive-listening studies on cerebral processing of emotional prosody, it supplements important findings on the network involved in perceiving emotional prosody as investigated during numerous active task-paradigms (e.g. Wildgruber et al., 2002, 2004 and 2005; Kotz et al., 2003; Mitchell et al., 2003). In addition to that, the study extends the investigated emotional intonations that elicit stronger hemodynamic responses than neutral prosody in right and left temporal cortices. Correlations of physical parameters with the hemodynamic responses of our region of interest showed that no single parameter alone can explain the stronger BOLD-signal in this area during perception of emotional stimuli. However, a multiple regression analysis with all five acoustic cues in combination revealed that the composition of all five parameters was sufficient to explain the increased responsiveness during comprehension of emotional prosody. Similarily, the perceived emotional arousal as one single parameter was effectual to predict increased responses within the right mid STG.

This result extends our understanding of the right mid STG as a voice-sensitive region being in particular sensitive to the interplay and therefore the more general percept which can be measured by the emotional arousal the single acoustic parameters create altogether. The investigation of these stimulus-driven effects formed the first part of my thesis.

The second part attended to the question whether the right mid STG is also sensitive to signals that are more related to the behavioural relevance expressed by emotional prosody and cannot be explained by differences of acoustic parameters alone. Erotically spoken stimuli are an excellent instrument to investigate such effects since in heterosexuals they might carry a direct signal for possible reproduction. Importantly, the same physical stimulus spoken in seducting tone of voice conveys different behaviourally relevant signals depending on the gender of speaker and listener (e.g. the prospect of a potential partner for reproduction versus a behaviourally less relevant utterance of a speaker of the same sex). In order to examine such cross-gender effects arousal ratings and responses of voice processing modules to seductive utterances spoken by individuals of different versus same sex as the listener were compared. Both female and male heterosexuals rated erotically spoken stimuli as higher arousing and showed stronger BOLD-signals in the right mid STG if they were expressed by actors of opposite versus same gender. This effect was restricted to erotic prosody and did not occur for any other emotional intonation investigated here.

These two different approaches further enlightened the different roles of the right mid STG in social communication: It does not only detect modulation within acoustic parameters but is also qualified to extract and decode additional relevant information that is not implemented in the physical stimulus as such but depends on contextual differences like listening to erotic voices of the same or of the opposite gender.

However, there are still open questions in the field of cerebral processing of emotion conveyed by the voice which should be addressed in future studies:

One open question is whether there are non-linear dependencies between acoustic parameters and the hemodynamic responses in right mid STG missed by our analytic approaches.

Moreover, in our study the emotional intensity (i.e. the absolute value of valence) was strongly correlated with arousal since only high-arousing emotional intonations were used. Future experiments should also include low-arousing emotions (e.g. like sadness) and contrast them with high-arousing ones (e.g. like fear) to better disentangle the effects due to different emotional categories from the impact of the varying emotional arousal.

Another important factor concerns the artificiality of our stimulus material since the emotions were posed by professional actors. Can one detect similar or even stronger effects with a more natural stimulus set (e.g. with intentional voices accidentally taken out of natural every-day situations)?
Finally, these remaining questions should be (and partly already are) addressed in future research experiments on emotional prosody and the brain.

5 Abschliessende Zusammenfassung

Die vorliegende Kernspintomografie-Studie zu emotionaler Prosodie wurde zum besseren Verständnis des Einflusses fünf verschiedener akustischer Parameter auf bestimmte stimmsensitive Areale innerhalb des menschlichen auditorischen Kortex durchgeführt.
Darüber hinaus sollte sie klären, ob akustische Parameter alleine für die stärkere Antwort stimmsensitiver Gehirnareale auf emotionale Prosodie (verglichen mit neutraler Sprachmelodie) verantwortlich sind oder ob es über ihren Einfluss hinaus mögliche Wirkungen verhaltensrelevanter Informationen gibt.
Als eine der wenigen Studien zur zerebralen Verarbeitung emotionaler Prosodie, die ohne aktive Aufgabe der Probanden durchgeführt wurde, kann sie frühere Befunde zu den an ihr beteiligten Arealen sinnvoll ergänzen und bestätigen (e.g. Wildgruber et al., 2002, 2004 und 2005; Kotz et al., 2003; Mitchell et al., 2003). Darüber hinaus erweiterte die Studie das Spektrum untersuchter emotionaler Intonationen, für die eine erhöhte Antwort stimmsensitiver Areale bereits nachgewiesen war, um die Kategorien “Erotik” und “Angst” und lässt so auf eine generalisierte Rolle des STG bei der Verarbeitung emotionaler Prosodie schliessen.
Innerhalb des rechten mittleren STG konnte das vorliegende Experiment den Einfluss der einzelnen akustischen Parameter auf die erhöhte hämodynamische Antwort während der Perzeption emotionaler Prosodie klären. Separate Regressionsanalysen zeigten, dass kein einzelner Parameter die erhöhten

Antworten ausreichend erklären konnte, das Zusammenspiel aller fünf analysierten Parameter jedoch sehr wohl. Die gleiche Beobachtung trifft auf das von den Probanden in einer Vorstudie eingeschätzte emotionale Arousal zu: Auch dieser Regressor konnte den Effekt verstärkter hämodynamischer Antworten unserer Zielregion während emotionaler Reize voraussagen.

Diese Resultate bestätigen und erweitern unser Verständnis des rechten mittleren STG als eine stimmsensitive Region, die vor allem für das Zusammenspiel mehrerer akustischer Parameter und weniger für einen speziellen Parameter sensibel ist. Plausibel ist in diesem Zusammenhang auch, dass sowohl alle fünf Parameter zusammen, wie auch das zuvor subjektiv eingeschätzte Arousal – welches letztlich ja auch durch die Modulation des akustischen Signals transportiert wird - höhere Aktivierungen während emotionalen Stimuli im rechten mittleren STG erklären konnten. Im ersten Teil der Doktorarbeit ging es um die oben beschriebene Untersuchung dieser stimulusgetriebenen Effekte.

Der zweite Teil sollte klären, ob der mittlere STG darüber hinaus ebenfalls für die Dekodierung von Informationen verantwortlich ist, die ihre Bedeutung für den Zuhörer durch die Geschlechterkonstellation der Interaktionspartner erhält.

Zur Untersuchung dieser Fragestellung stellen erotisch eingesprochene Reize ein exzellentes Mittel dar, da sie innerhalb der Kommunikation heterosexueller Individuen direkte fortpflanzungsrelevante Informationen tragen können. Der physikalisch-identische Stimulus mit erotischer Intention kann abhängig vom Geschlecht des Sprechers und Zuhörers einen völlig anderen Bedeutungszuwachs bekommen (z. B. ein Signal eines möglichen Partners oder eine weitaus weniger aufregende Äußerung eines Gleichgeschlechtlichen). Erotisch-eingesprochene Reize können somit ausgezeichnet dazu verwendet werden, über physikalische Parameter hinausgehende Veränderungen und deren Auswirkungen auf Verhalten und Gehirnantwort Aufschluss zu geben. Um diese Auswirkungen genauer zu betrachten, wurden sowohl behaviorale Arousalratings als auch hämodynamische Signale des rechten mittleren STG auf erotisch eingesprochene Stimuli gleichgeschlechtlicher Sprecher mit Situationen verglichen, in denen Sprecher und Zuhörer unterschiedlichen Geschlechts waren.

Interessanterweise zeigten die Resultate nicht nur höhere Arousalratings auf Verhaltensebene bei der Perzeption erotischer Prosodie, sondern ebenfalls erhöhte hämodynamische Antworten im mittleren rechten STG, wenn die erotisch gesprochenenen Stimuli von einem gegengeschlechtlichen Sprecher stammten. Dieser Effekt zeigte sich exklusiv für erotische Prosodie und konnte für keine der anderen emotionalen Kategorien nachgewiesen werden.

Die beiden in dieser Doktorarbeit verwendeten Ansätze konnten also zu einem besseren Verständnis des rechten mittleren STG in sozialer Kommunikation beitragen: Der rechte mittlere superiore temporale Gyrus scheint nicht nur sensitiv für die Modulation prosodisch relevanter akustischer Parameter, sondern ebenfalls dazu befähigt, darüber hinaus gehende Informationen hoher biologischer Relevanz zu dekodieren.

Trotz dieser neuen Erkenntnisse hat die vorliegende Studie ebenfalls einige zu klärende Fragen offen gelassen, die in zukünftigen Studien untersucht werden sollten:

So wäre eine Klärung sinnvoll, ob und in welchem Ausmass nicht-lineare Korrelationen zwischen akustischen Parametern und hämodynamischem Verhalten im rechten mittleren STG existieren.

Darüber hinaus wurden in unserer Stimulusauswahl nur emotionale Stimuli mit hohem Arousal berücksichtigt. In einer zukünftigen Studie sollten explizit emotionale Intonationen mit geringem emotionalen Arousal (z. B. Trauer) und deren Einfluss auf den rechten mittleren STG, im Vergleich mit Emotionen hohen Arousals (z. B. Angst) untersucht werden, um den Effekt der Emotion an sich besser von dem des vermittelten Arousals trennen zu können.

Schliesslich stellt sich die Frage, ob sich ähnliche oder sogar stärkere Effekte mit einem natürlicheren Set an Stimulusmaterial (beispielsweise aus natürlichen Alltagssituationen entnommene Beispiele emotionaler Prosodie) reproduzieren lassen.

6 References

Alter, K., Rank, E., Kotz, S.A., Toepel, U., Besson, M., Schirmer, A., Friederici, A.D., 2003. Affective encoding in the speech signal and in event-related potentials. Speech Commun. 40, 61–70.

Andersson, J.L., Hutton, C., Ashburner, J., Turner, R., Friston, K.J., 2001. Modeling geometric deformations in EPI time series. NeuroImage 13, 903–919.

Banse, R., Scherer, K., 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. 70, 614–636.

Barrett, D.J.K., Hall, D.A., 2006. Response preferences for “what” and “where” in human non-primary auditory cortex.NeuroImage 32, 968–977.

Beaucousin, V., Lacheret, A., Turbelin, M.R., Morel, M., Mazoyer, B., Tzourio-Mazoyer, N., 2007. FMRI study of emotional speech comprehension. Cereb. Cortex 17, 339–352.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Boersma, P., 2001. Praat, a system for doing phonetics by computer. Glot Int. 5, 341–345.

Boul, K.E., Funk, W.C., Darst, C.R., Cannatella, D.C., Ryan, M.J., 2007. Sexual selection drives speciation in an Amazonian frog. Proc. Biol. Sci., 274, 399–406.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25, 49–50.

Buxton, R.B., Uludag, K., Dubowitz, D.J., Liu, T.T., 2004. Modelling the hemodynamic response to brain activation. NeuroImage 23 (Suppl. 1), 220–233.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18, 192–205.

Costa, M., Braun, C., Birbaumer, N., 2003. Gender differences in response to pictures of nudes: a magnetoencephalographic study. Biol. Psychol., 63, 129–147.

Crinion, J.T., Lambon-Ralph, M.A., Warburton, E.A., Howard, D., Wise, R.J.S., 2003. Temporal lobe regions engaged during normal speech comprehension. Brain 126, 1193–1201.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. 18, 193–222.

Eickhoff, S., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25, 1325–1335.

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. 572–595.

Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., Grodd, W., Wildgruber, D., 2006a. Effects of prosodic emotional intensity on activation of associative auditory cortex. NeuroReport 17, 249–253.

Ethofer, T., Anders, S., Erb, M., Herbert, C.,Wiethoff, S., Kissler, J., Grodd, W., Wildgruber, D., 2006b. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. NeuroImage 30, 580–587.

Ethofer, T., Wiethoff, S., Anders S., Kreifelts, B., Grodd, W., Wildgruber D., 2007. The voices of seduction: cross-gender effects in processing of erotic prosody. SCAN 2, 2334-2337.

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. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Glaser, D., Henson, R., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. NeuroImage 16, 484–512.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R.,Vuilleumier, P., 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. Nat. Neurosci. 8, 145–146.

Griffiths, T.D., 2003. Functional imaging of pitch analysis. Ann. N. Y. Acad. Sci. 999, 40–49.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W., 1999. “Sparse” temporal sampling in auditory fMRI. Hum. Brain Mapp. 7, 213–223.

Hauser, M.D., 1993. Rhesus monkey copulation calls: honest signals for female choice? Proc. Biol. Sci., 254, 93–96.

Henson, R.N.A., 2003. Neuroimaging studies of priming. Prog. Neurobiol. 70, 53–81.

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.

Ishai, A., Pessoa, L., Bikle, P.C., Ungerleider, L.G., 2004. Repetition suppression of faces is modulated by emotion. Proc. Natl. Acad. Sci. 29, 9827–9832.

Jäncke, L., Shah, N.J., Posse, S., Grosse-Ryuken, M., Müller-Gärtner, H.W., 1998. Intensity coding of auditory stimuli: an fMRI study. Neuropsychologia 36, 875–883.

Jäncke, L., Wüstenberg, T., Scheich, H., Heinze, H.J., 2002. Phonetic perception and the temporal cortex. NeuroImage 15, 733–746.

Joanisse, M.F, Gati, J.S., 2003. Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. NeuroImage 19, 64–79.

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. Hum. Brain Mapp., 16, 1–13.

Kiflawi, M., Gray, D.A., 2000. Size dependent response to conspecific mating calls by male crickets. Proc. Biol. Sci., 267, 2157–2161.

Kotz S. A., Meyer M., Alter K., Besson M., von Cramon D. Y., Friederici A. D., 2003. On the lateralization of emotional prosody: an event-related functional MR investigation. Brain Lang. 86: 366-376.

Langers, D.R.M., van Dijk, P., Schoenmaker, E.S., Backes, W.H., 2007. fMRI activation in relation to sound intensity and loudness. NeuroImage 35, 709–718.

Lattner, S., Meyer, M.E., Friederici, A.D., 2005. Voice perception: sex, pitch, and the right hemisphere. Hum. Brain Mapp. 24, 11–20.

LeDoux, J.E., 1993. Emotional memory: in search of systems and synapses. Ann. N. Y. Acad. Sci. 702, 149–157.

London, M., Hausser, M., 2005. Dendritic computation. Annu. Rev. Neurosci. 28, 503–532.

Mathiak, K., Rapp, A., Kircher, T.T.J., Grodd,W., Hertrich, I.,Weiskopf, N., Lutzenberger, W., Ackermann, H., 2002. Mismatch responses to randomized gradient switching noise as reflected by fMRI and wholehead magnetoencephalography. Hum. Brain Mapp. 16, 190–195.

Mehrabian, A, 1972. Nonverbal communication, Aldine-Atherton, Chicago, Illinois.

Meyer, M., Zaehle, T., Gountouna, V.-E., Barron, A., Jäncke, L., Turk, A., 2005. Spectro-temporal processing during speech perception involves left posterior auditory cortex. NeuroReport 16, 1985–1989.

Mitchell R. L., Elliott R., Barry M., Cruttenden A., Woodruff P. W., 2003. The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. Neuropsychologia. 41, 1410-1421.

Mohnrad-Kohn, GH, 1947. The prosodic qualities of speech and its disorders a brief survex from a neurologist´s point of view. Acta Psychiatr Neurol Scand 22: 255-269.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. NeuroImage 13, 684–701.

Mottonen, R., Calvert, G.A., Jääskeläinen, I.P., Matthews, P.M., Thesen, T., Tuomainen, J., Sams, M., 2006. Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. NeuroImage 30, 563–569.

Näätänen, R., 1990. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. Behav. Brain Sci. 13, 201–288.

Oldfield, R.C., 1971. Assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Olson, I.R., Plotzker, A., Ezzyat, Y., 2007. The enigmatic temporal pole: a review of findings on social and emotional processing. Brain 130, 1718–1731.

Pittam J & Scherer KR, 1993. Vocal expression and communication of emotion. In Lewis M/Haviland JM (Eds.) Handbook of emotions, New York: Guilford Press, pp. 185-198.

Purdon, P.L.,Weisskoff, R.M., 1998. Effect of temporal autocorrelations due to physiological noise stimulus paradigm on voxel-level false positive rates in fMRI. Hum. Brain Mapp. 6, 239–249.

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H.J., Zilles, K., 2001. Probabilistic mapping and volume measurement of human primary auditory cortex. NeuroImage 13, 669–683.

Rauschecker, J.P., Tian, B., 2000. Processing of “what” and “where” in auditory cortex. Proc Natl. Acad. Sci. 97, 11800–11806.

Reiterer, S.M., Erb, M., Droll, C.D., Anders, S., Ethofer, T., Grodd, W., Wildgruber, D., 2005. Impact of task difficulty on lateralization of pitch and duration discrimination. NeuroReport 16, 239–242.

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.

Schall, U., Johnston, P., Todd, J., Ward, P.B., Michie, P.T., 2003. Functional neuroanatomy of auditory mismatch processing: an event-related fMRI study of duration-deviant oddballs. NeuroImage 20, 729–736.

Scherer, K.R., 1986. Vocal affect expression: a review and a model for future research. Psychol. Bull. 99, 143–165.

Scherer, K.R., 2003. Vocal communication of emotion: a review of research paradigms. Speech Commun. 40, 227–256.

Schmidt, C.F., Zaehle, T., Meyer, M., Geiser, E., Boesinger, P., Jäncke, L.,in press. Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. Hum. Brain Mapp. 29: 46-56.

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.

Sokhi, D.S., Hunter, M.D., Wilkinson, I.D., Woddruff, P.W.R., 2005. Male and female voices activate distinct regions in the male brain. NeuroImage 27, 572–578.

Surguladze, A., Brammer, M.J., Young, A.W., et al., 2003. A preferential increase in the extrastriate response to signals of danger. NeuroImage, 19, 1317–1328.

Tervaniemi, M., Szameitat, A.J., Kruck, S., Schröger, E., Alter, K., De Baene, W., Friederici, A.D., 2006. From air oscillations to music and speech: functional magnetic resonance imaging evidence for fine-tuned neural networks in audition. J. Neurosci. 26, 8647–8652.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.

von Kriegstein, K.V., Giraud, A.L., 2004. Distinct functional substrates along the right superior temporal sulcus for the processing of voices. NeuroImage 22, 948–955.

von Kriegstein, K.V., Kleinschmidt, A., Sterzer, P., Giraud, A.L., 2005. Interaction of face and voice areas during speaker recognition. J. Cogn. Neurosci., 17, 367–376.

Vuilleumier, P., Schwartz, S., 2001. Emotional facial expressions capture attention. Neurology 56, 153–158.

Vuilleumier, P., 2005. How brains beware: neural mechanisms of emotional attention. Trends Cogn. Sci. 95, 85–94.

Warren, J.D., Jennings, A.R., Griffiths, T.D., 2005. Analysis of the spectral envelope of sounds by the human brain. NeuroImage 24, 1052–1057.

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.

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.

Wildgruber D., Pihan H., Ackermann H., Erb M., Grodd W., 2002. Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence and sex. NeuroImage 15: 856-869.

Wildgruber D., Hertrich I., Riecker A., Erb M., Anders S., Grodd W., Ackermann H., 2004. Distinct frontal regions subserve evaluation of linguistic and affective aspects of intonation. Cereb Cortex 14: 1384-1389.

Wildgruber D., Riecker A., Hertrich I., Erb M., Grodd W., Ethofer T., Ackermann H., 2005. Identification of emotional intonation evaluated by fMRI. NeuroImage 24:1233-1241.

Zaehle, T., Schmidt, C.F., Meyer, M., Baumann, S., Baltes, C., Boesinger, P., Jäncke, L., 2007. Comparison of “silent” clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. NeuroImage 37, 1195–1204.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46.

Zatorre, R.J., Bouffard, M., Belin, P., 2004. Sensitivity to auditory object features in human temporal neocortex. J. Neurosci. 24, 968–977.

7 Danksagung

Ich bedanke mich sehr herzlich für die äusserst fachgerechte und unkomplizierte Hilfe und Diskussionsfreudigkeit meines Doktorvaters PD Dr. med. Dirk Wildgruber, der mir stets mit Rat und guten Vorschlägen zur Seite stand.
Ausserdem danke ich meinem Betreuer Dr. med. Thomas Ethofer, der mir die Möglichkeit gegeben hat, mein Interesse an den Neurowissenschaften durch meine experimentelle Arbeit unter seiner stetigen Anleitung praktisch auszuleben und meinen neurowissenschaftlichen Blick weit über diese Dissertation hinweg zu schulen und zu erweitern. DANKE!
Bedanken möchte ich mich auch bei Prof. Dr. Grodd und den Mitarbeitern seiner Sektion (vor allem Michael Erb und Kurt Prangenberg), die mit den Apparaturen und fachlicher Hilfe einen enormen Anteil am Gelingen meiner Studie tragen.
Benjamin Kreifelts hat als Mitglied der Arbeitsgruppe an zahlreichen Stellen vielfältig und entscheidend geholfen, wofür ich mich herzlich bedanken möchte.
Hubertus Becker möchte ich für seinen Rat bei der Stimulusproduktion und Normalisierung dankend erwähnen.
Ausserdem sei Johanna Kissler und Cornelia Herbert für die semantische Evaluation der verwendeten Nomen und Adjektive gedankt.

8 Curriculum vitae

Wiethoff, Sarah

Personal	Date of birth: 13.08.85 Place of birth: Nordhorn, Germany
Education	10/2004 – 10/2010 Medical student at the Universities of Tübingen & Maribor (Slovenia) – Graduation: 22.10.2010 6/2004 High School Leaving Examination (Abitur) in Nordhorn

Selected Publications

Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T. Cerebral processing of emotional prosody-influence of acoustic parameters and arousal.
Neuroimage. 2008; 39:885-893.

Wiethoff S, Wildgruber D, Grodd W, Ethofer T. Response and habituation of the amygdala during processing of emotional prosody. Neuroreport. 2009; 20: 1356-1360.

Ethofer T, Kreifelts B, **Wiethoff S**, Wolf J, Grodd W, Vuilleumier P, Wildgruber D. Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. J Cogn Neurosci. 2009; 21: 1255-1268.

Ethofer T, **Wiethoff S**, Anders S, Kreifelts B, Grodd W, Wildgruber D. The voices of seduction: cross-gender effects in processing of erotic prosody. Soc Cogn Affect Neurosci. 2007; 2: 334-337.

Awards

2010 -	Carl-Liebermeister-Promotionspreis, University of Tübingen
2009 -	Hans-Heimann-Promotionspreis, Deutsche Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde
2004 - 2010	Studienstiftung des Deutschen Volkes (German National Merit Foundation)-Scholarship
2006 - 2007	Forum Scientiarum - Udo-Keller-Foundation Scholarship
2006 -	Travel Award of the Human Brain Mapping Organization

Volunteerism and experience abroad

I have been engaged in improving education, politics and the religious life of my hometown Nordhorn (Germany) for many years and was elected to represent students at my high school.
In Tübingen I headed a project involving inmates in athletic competitions and serve as Local Exchange Officer for the International Federation of Medical Students Association.

2001+2002	School term at Lycée Privée in Saumur (France)
04/2003	Visited a school in Kerang (Australia)
12/2003	Visited an African community in Ghana
07+08+09/2004	Internship with CAMPHILL – LE BEAL „Foyer de Vie“ (France)
03/2006	Language training in Salamanca (Spain)
08+09/2007	Radiology-Internship at Harvard Medical School, Boston (USA)
02+03/2008	Research Internship at the Department of Fundamental Neurosciences, Geneva (Switzerland)
08+09/2008	Children-surgery Internship in Lima (Peru)
2008+2009	Universitary term at the University of Maribor (Slovenia)

Leisure activities, sports

I like playing the piano and juggling. In sports I am especially fond of triathlon and marathon (runner up at the German Universitary Championships 2009 and 2010).

9 Publikationsliste

Wiethoff S, Wildgruber D, Grodd W, Ethofer T (2009). Response and habituation of the amygdala during processing of emotional prosody. Neuroreport. 2009; 20: 1356-1360.

Ethofer T, Kreifelts B, **Wiethoff S**, Wolf J, Grodd W, Vuilleumier P, Wildgruber D (2009). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. J Cogn Neurosci. 2009; 21: 1255-1268.

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. 2008; 39:885-893.

Wildgruber D, **Wiethoff S**, Anders S, Kreifelts B, Grodd W, Ethofer T (2007). Stimmen der Verführung: Geschlechtsspezifische Effekte bei der zerebralen Verarbeitung erotischer Sprachmelodie. Kongress der Deutschen Gesellschaft für Psychiatrie, Psychotherapie und Nervenheilkunde, Berlin, Der Nervenarzt. 2007; 78: Supplement 2, S. 251.

Ethofer T, **Wiethoff S**, Anders S, Kreifelts B, Grodd W, Wildgruber D (2007). The voices of seduction: cross-gender effects in processing of erotic prosody. Soc Cogn Affect Neurosci. 2007; 2: 334-337.

Ethofer T, **Wiethoff S**, Kreifelts B, Herbert C, Grodd W, Wildgruber D (2007). Processing of emotional prosody – Influence of acoustic parameters. 13th Int. Conference on Functional Mapping of the Human Brain, Chicago, NeuroImage 27: Supplement on CD-ROM.

Wiethoff S, Wildgruber D, Becker H, Herbert C, Erb M, Grodd W, Ethofer T (2006) Perception of emotional information in the voice: an fMRI-study. 12th Int. Conference On Functional Mapping of the Human Brain, Florence, NeuroImage 26: Supplement on CD-ROM.

Ethofer T, Anders S, **Wiethoff S**, Erb M, Herbert C, Saur R, Grodd W, Wildgruber D (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. 12th Int. Conference on Functional Mapping of the Human Brain. Florence, NeuroImage 26: Supplement on CD-ROM.

Ethofer T, Anders S, **Wiethoff S**, Erb M, Herbert C, Saur R, Grodd W, Wildgruber D (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. Neuroreport. 2006; 17:249-253.

Ethofer T, Anders S, Erb M, Herbert C, **Wiethoff S**, Kissler J, Grodd W, Wildgruber D (2006). Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. NeuroImage. 2006; 30:580-587.

Wiethoff S, Baier N, Paret C (2008). Zeit und Zeitbewusstsein, in “Kognition und Verhalten”. LIT Verlag. 2008, Hg.: Dirk Evers, Niels Weidtmann.