

## Research report

# Recognition of emotional prosody and verbal components of spoken language: an fMRI study

Tony W. Buchanan<sup>a,b,\*</sup>, Kai Lutz<sup>b,c</sup>, Shahram Mirzazade<sup>b</sup>, Karsten Specht<sup>b</sup>, N. Jon Shah<sup>b</sup>,  
Karl Zilles<sup>b</sup>, Lutz Jäncke<sup>b,c,1</sup>

<sup>a</sup> Department of Psychiatry and Behavioral Sciences, University of Oklahoma Health Sciences Center, and Veterans Affairs Medical Center, Oklahoma City, OK, USA

<sup>b</sup> Institute of Medicine, Research Center Jülich, Jülich, Germany

<sup>c</sup> Institute of General Psychology, Otto-von-Guericke University, Magdeburg, Lennéstraße 6, D-39112, Magdeburg, Germany

Accepted 23 November 1999

---

**Abstract**

This study examined the neural areas involved in the recognition of both emotional prosody and phonemic components of words expressed in spoken language using echo-planar, functional magnetic resonance imaging (fMRI). Ten right-handed males were asked to discriminate words based on either expressed emotional tone (angry, happy, sad, or neutral) or phonemic characteristics, specifically, initial consonant sound (bower, dower, power, or tower). Significant bilateral activity was observed in the detection of both emotional and verbal aspects of language when compared to baseline activity. We found that the detection of emotion compared with verbal detection resulted in significant activity in the right inferior frontal lobe. Conversely, the detection of verbal stimuli compared with the detection of emotion activated left inferior frontal lobe regions most significantly. Specific analysis of the anterior auditory cortex revealed increased right hemisphere activity during the detection of emotion compared to activity during verbal detection. These findings illustrate bilateral involvement in the detection of emotion in language while concomitantly showing significantly lateralized activity in both emotional and verbal detection, in both the temporal and frontal lobes. © 2000 Elsevier Science B.V. All rights reserved.

**Keywords:** Emotion; Language; Prosody; fMRI

---

**1. Introduction**

Spoken language is able to convey a wealth of information beyond the obvious linguistic meaning of words. The term prosody describes the non-propositional cues that may be passed along through language, including intonation, stresses, and accents [47,49]. Emotional prosody involves the expression of emotion through the intonation of

spoken language. A great deal of work in neuropsychology and behavioral neurology has focused on prosody, or the lack thereof, in patients with brain damage [see Ref. [47]]. The term ‘aprosodia’ has been used to describe deficits either in the expression or understanding of prosody. Specifically, Ross [47] describes lesions in the right hemisphere resulting in a pattern of aprosodias analogous to the well-documented pattern of left hemisphere lesions resulting in the various aphasia.

The anatomy and physiology of the auditory cortex has been studied extensively in both humans [4,18,22,49,53,54] and primates [19,21,28,44]. Language-related functions of the brain have been studied extensively since the advent of functional neuroimaging [2–6,38,39]. This work has been seminal in understanding the neuroanatomy of language as well as in establishing the methodology of functional neuroimaging [see Ref. [5] for review]. Imaging studies of emotion have begun to provide a picture of the neural

---

\* Corresponding author. Behavioral Sciences Laboratory, University of Oklahoma Health Sciences Center, Veterans Affairs Medical Center (151A), 921 Northeast 13th Street, Oklahoma City, OK 73104, USA. Fax: +1-405-297-5909; e-mail: tony@ouhsc.edu

<sup>1</sup> Also corresponding author. Tel.: +49-391-6714842; fax: +49-391-6714815; e-mail: lutz.jaencke@gse-w.uni-magdeburg.de.

activity during both the experience and the detection of emotion [1,30–33,45]. Only a few studies, however, have examined the expression of emotion in language using functional neuroimaging [20,23,40,55].

Two studies using positron emission tomography (PET) have examined the recognition of emotional prosody in spoken language [20,23]. George et al. [20] documented increased activity in the right prefrontal cortex during detection of emotional prosody expressed in sentences. Additionally, these authors reported significant left prefrontal activation when volunteers were instructed to listen for content of verbal sentences compared with listening for prosody. Imaizumi et al. [23] compared the detection of emotional prosody with the detection of speakers differing in age and gender. These authors also documented significant right prefrontal activity during the detection of emotional prosody.

Neither the aforementioned clinical studies nor neuroimaging work have established a clear picture of the neural substrates of emotional prosody. Differences in methodology including different imaging techniques, different verbal stimuli and task instructions may be to blame for these disparate findings. For example, in the two aforementioned PET studies [20,23], one involved the detection of emotional prosody within a sentence and required a verbal response [20] while the other involved detection of prosody expressed in a single word compared to the detection of speakers of different age and gender [23].

The rationale for the current study was to examine the neural regions involved in the processing of emotional prosody compared with the processing of phonemic discrimination. In order to control for potential confounds arising from the use of different stimuli and response parameters each condition consisted of the presentation of the exact same stimuli (in different, random orders), differing only in task instructions. In one task, the volunteers were asked to detect the emotional intonation of spoken words while in the other, they were asked to merely detect a target word irrespective of emotional tone. Comparison between the different detection conditions permitted identification of the neural activity associated with the detection of both emotional prosody and verbal components of language. We predicted that the detection of emotion would result in more right hemispheric activity and that detection of language would result in more left hemispheric activity, and specifically that these differences would be most evident in the auditory-related cortex.

## 2. Materials and methods

### 2.1. Subject characteristics

Ten male right-handed volunteers (tested with a standard hand preference test [37]), ranging in age from 22 to

40 years, with no history of neurological or audiological illness were studied. After a full explanation of the nature and risks of the research, subjects gave informed written consent for all studies according to a protocol approved by the Ethics Committee of the Heinrich-Heine University, Düsseldorf. Eight of the ten subjects were native speakers of German (one was a native English speaker and the other a native speaker of Turkish). All subjects spoke English and understood the meaning of the four words.

### 2.2. Stimulus materials and experimental conditions

The auditory stimuli consisted of the four words, 'bower', 'dower', 'power', and 'tower', each originally spoken by a native English-speaking male phonetician in angry, happy, neutral, and sad tones of voices. Among the four words, the phonetic pronunciations 'bower' and 'dower' have meaning in both English and German (although they are spelled differently in German). The other two words, 'power' and 'tower,' although they are not part of the standard German lexicon, are widely used in the German language and known to all the subjects who participated in this study. These stimuli have been used previously in dichotic listening studies examining verbal and affective laterality effects and have been shown to have adequate discriminative properties [7,8]. This experiment involved only the binaural presentation of these stimuli. The stimuli were edited to a common length of approximately 650 ms and equalized in intensity. They were then stored in 16-bit, digital format on a personal computer.

Each subject was allowed to listen to the sounds of all the stimuli as often as he liked prior to entering the scanner in order to become familiar with both the verbal and affective characteristics of the stimuli. The stimuli consisted of each of the four words (bower, dower, power, and tower) spoken in each of the four emotions (angry, happy, neutral, and sad) resulting in 16 separate word/emotion combinations. In each of the four experimental conditions, subjects were asked to listen for a different target and press a button when that target was detected. Each condition involved instructions to listen for a different target: (1) 'Power' spoken in any of the four emotional tones (phonetic task), (2) 'Bower' spoken in any of the four emotional tones (phonetic task), (3) Any of the four words spoken in a 'Happy' tone (emotional task), and (4) Any of the four words spoken in a 'Sad' tone (emotional task). In each condition, 108 stimuli were presented consisting of a random mix of all 16 separate word/emotion combinations, including 56 targets in each. The order of target stimulus condition as well as the order of stimulus presentation within each condition was randomized for each subject. The interval between words was 1 s. Performance data were only recorded from nine of the ten subjects due to technical difficulties.

### 2.3. Procedure

Functional MR images were acquired using a 1.5 Tesla Siemens MRI system (SIEMENS Magnetom Vision, Erlangen, FRG), equipped with echo planar imaging (EPI) capability and a standard radiofrequency (RF) head coil for transmit and receive. Pulse sequence parameters were as follows: gradient echo EPI; repetition time (TR) = 6 s; echo time (TE) = 66 ms; field of view (FOV) =  $200 \times 200$  mm; flip angle ( $\alpha$ ) =  $90^\circ$ ; matrix size =  $64 \times 64$ ; in-plane resolution =  $3.125 \times 3.125$  mm; slice thickness = 5.0 mm; inter-slice gap = 0.3 mm. Using a mid-sagittal scout image, 16 axial slices were oriented in the anterior–posterior commissure (AC-PC) plane, with the lowermost slice positioned to be 20 mm below the AC-PC line. In addition, high-resolution, T1-weighted anatomical images of the entire brain were obtained in 3D using the mp-rage (magnetization-prepared, rapid acquisition gradient echo) pulse sequence with the following parameters: TR = 11.4 ms; TE = 4.4 ms; TI = 300 ms;  $\alpha = 15^\circ$ ; 1 excitation; FOV = 230 mm; matrix =  $200 \times 256$ ; 128 sagittal slices with 1.25 mm slice thickness.

During scanning the room lights were dimmed and the subjects' eyes were open. Auditory stimuli were presented binaurally using a digital playback system and a magnetically shielded transducer system. The acoustic stimulation system terminated in tightly occlusive headphones allowing unimpeded conduction of the stimulus with good suppression of ambient scanner noise by about 20 dB. During each experimental condition, a series of 78 data sets were acquired. Each series consisted of multiple periods of 'baseline' (OFF), during which, subjects heard only the ambient machine noise, alternating with periods of 'activation' (ON), during which prepared auditory stimuli were delivered. Each series began with six baseline data sets (36-s interval), followed by 72 images during which 'rest' alternated with 'activation' every 54 s (1 series = 4 cycles, 108 s/cycle, 18 images/cycle). The total duration of each image series was about 8 min.

Because of the limitations involved in taking an SPL meter into an MRI scanner, the intensity of the stimuli (mean signal intensity (SPL) during a 30 s epoch) was determined outside the scanner using an artificial ear (Bruel and Kjaer KA637) wearing the headphones. The acoustic noise level within the MRI scanner during an EPI measurement was measured 1 m from the bore of the scanner magnet with a capacitance microphone revealing a root mean square level of 90 to 100 dB. Because the attenuation factor of the headphones was about 20 dB, the average intensity of the scanner noise perceived by the subjects was estimated to be about 70 to 80 dB. Frequency analysis of the noise generated by the EPI sequence recorded in the scanner revealed five prominent frequencies. The corresponding amplitudes relative to the maximum amplitude were: 612 Hz (−15 dB), 1847 Hz (−29 dB), 3036 Hz (−21 dB), 3365 Hz (−37 dB), and 3644 Hz (−36 dB).

This analysis revealed no overlap of the prominent frequencies of the EPI-noise with the formant frequencies of the diphthong/au/contained within all stimulus words. The diphthong/au/is generated with downward formant transitions for both formants. The beginning of the formant F1 was measured to lie on average at 800 Hz and at the end at 400 Hz. For F2, the beginning was measured to be 1300 Hz and 700 Hz at the end. In order to reduce possible masking effects of scanner noise on the experimental stimuli, a TR of 6 s was adopted [see Ref. [52]]. Stimulus presentation was triggered by the EPI sequence to fall in between the interscan gap of 4 s (sequence scanning time = 2 s) resulting in a non-masked presentation of the stimuli (3 stimuli per sequence).

### 2.4. Image analysis

Image analysis was performed on an Ultra 4 workstation (Sun Microsystems) using MATLAB (Mathworks, Natick, MA, USA) and SPM97 software (SPM software, MRC Cyclotron Unit, London) [17]. The first six images of each time-series, during which the MR signal reaches a steady-state, were discarded. The 72 remaining volume images of each condition were automatically realigned to the first image to correct for head movement between scans [15]. The images of the four conditions were then co-registered and spatially normalized into the Montreal Neurological Institute (MNI) template [11]. The procedure starts with a 12-parameter affine transformation and a 6-parameter three-dimensional quadratic (or second order) deformation followed by non-linear (plastic) deformations on a slice by slice basis using Fourier-like basis functions; the parameters are estimated using a least-squares approach after linearizing the problem [15]. In this space, one pixel represents  $4 \times 4$  mm in the  $x$  and  $y$  dimensions, with an interplanar distance of 4 mm. These spatially transformed functional data sets from each subject were smoothed slightly with a Gaussian filter of root-mean-square radius of 8 mm to improve the signal-to-noise ratio, approximate normal distribution of the data, and correct for inter-subject variability of sulcal/gyral anatomy. Voxels that had values greater than 0.8 of the volume mean in all the images were selected to restrict analysis to intracranial regions. The effects of global activity (whole volume) and time were removed using linear regression and sine/cosine functions as confounds (up to a maximum of two cycles per 108 scans). Removing the latter confounds corresponds to high-pass filtering the time series to remove low frequency artifacts due to cardio-respiratory and other cyclical components.

### 2.5. Statistical parameter mapping

The spatially normalized fMRI time-series data of each subject were analyzed separately in order to closely exam-

ine each individual within the framework of the Talairach co-ordinates [56]. Significantly activated pixels were searched for by using the ‘General Linear Model’ approach for time-series data suggested by Friston and colleagues [13–16]. An ANOVA was calculated for each voxel of the functional data set comparing the four conditions to rest. In addition, we calculated interactions in order to detect between-condition differences (i.e., verbal vs. emotion). The resulting set of voxel values for each contrast constitutes a statistical parametric map of the  $t$  statistic ( $SPM\{t\}$ ). The  $SPM\{t\}$  were transformed to the unit normal distribution  $SPM\{Z\}$ . Voxels were identified as significant if they passed the height threshold of  $z = 3.09$  ( $p = 0.001$ , uncorrected) and a spatial extent threshold of  $k = 20$  ( $p = 0.05$ , corrected for multiple comparisons).

The resulting foci were then characterized in terms of peak height ( $u$ ) and spatial extent ( $k$ ). The significance of each region was estimated using the probability that the peak height observed could have occurred by chance [ $p(Z_{\max} > u)$ ] or that the observed number of voxels could have occurred by chance [ $p(n_{\max} > k)$ ] over the entire volume analyzed [16]. The activated voxels surviving this procedure were superimposed on ‘SPM brain projections’ and on individual spatially normalized high-resolution MR-anatomical scans.

The spatially transformed functional data sets were averaged across subjects. As in the individual analyses only pixels passing a height threshold of  $z = 3.09$  and a cluster threshold of  $p < 0.05$  corrected for multiple comparisons were considered significant. The activated voxels surviving this procedure were superimposed on ‘SPM brain projections’.

Fig. 1 shows the defined regions of interest (ROI) comprising the anterior and posterior auditory cortices in order to conduct planned comparisons on activity in these areas in both the right and left hemisphere. The rationale for the focus on the auditory cortex stems from the authors’ interest in both the structure and function of these areas as well as by studies documenting differences in auditory cortex activity during different types of auditory tasks. These ROIs were defined according to the Talairach atlas [56] and prominent sulcal and gyral landmarks (Heschl’s gyrus, sylvian fissure, ramus posterior ascendens, ramus posterior descendens, and superior temporal sulcus [18,24,53,54]) to include: (1) the primary and immediate auditory area and (2) the auditory association area. The exact spatial delineation of the auditory cortices is currently unknown for the following reasons: (i) the spatial extent of the cytoarchitectonic area of the primary auditory cortex is currently unknown, (ii) the exact delineation of the anatomical landmarks is not possible because of the limited spatial resolution of the normalized fMRI images ( $4 \text{ mm}^3$ ), (iii) the angulated shape of the Heschl’s gyri and the Sylvian fissure vary considerably in horizontal and vertical directions within subjects (left vs. right hemisphere) and between subjects. This variability prevents a

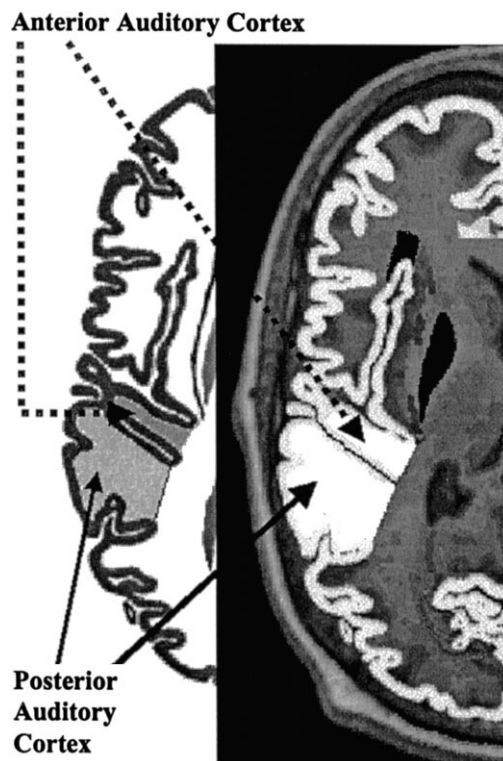


Fig. 1. Map of the regions of interest (ROIs) superimposed on transverse views of the brain. The exact borders of the anterior auditory cortices were defined on Talairach coordinates (left anterior auditory cortex:  $-65 < x < -28$ ,  $-36 < y < -8$ ,  $8 < z < 16$ ; right anterior auditory cortex:  $28 < x < 65$ ,  $-36 < y < -8$ ,  $8 < z < 16$ ). The exact borders of the posterior auditory cortices were defined on Talairach coordinates (left posterior auditory cortex:  $-70 < x < -28$ ,  $-50 < y < -35$ ,  $4 < z < 20$ ; right posterior auditory cortex:  $28 < x < 70$ ,  $-50 < y < -35$ ,  $4 < z < 20$ ).

simple right vs. left comparison of the auditory cortices using the standard SPM procedures. Additionally, anterior and posterior auditory cortex activity cannot be quantitatively analyzed using these procedures. Thus, we defined the anterior auditory cortex as a region covering the Heschl’s gyrus defined by Heschl’s sulcus as the anterior border and the first transverse sulcus as the posterior border. The exact borders of the anterior auditory cortices were defined on Talairach coordinates (left anterior auditory cortex:  $-65 < x < -28$ ,  $-36 < y < -8$ ,  $8 < z < 16$ ; right anterior auditory cortex:  $28 < x < 65$ ,  $-36 < y < -8$ ,  $8 < z < 16$ ). The posterior auditory cortex was defined as the area immediately posterior to the anterior auditory cortex region with the first transverse sulcus as the anterior border and the bifurcation of the Sylvian fissure as the posterior border. The exact borders of the posterior auditory cortices were defined on Talairach coordinates (left posterior auditory cortex:  $-70 < x < -28$ ,  $-50 < y < -35$ ,  $4 < z < 20$ ; right posterior auditory cortex:  $28 < x < 70$ ,  $-50 < y < -35$ ,  $4 < z < 20$ ). For these ROIs, numbers of significantly activated voxels, as well as

mean signal intensity changes relative to baseline were determined for each subject individually. The number of activated voxels in both the anterior and posterior auditory cortices for each hemisphere were entered into a two Task (Emotional Detection and Verbal Detection) by two Hemisphere (Right and Left) repeated measures, within subjects multivariate analysis of variance (MANOVA). Post-hoc analyses were then conducted between conditions (detect bower/power/happy/sad) within each task for activity within each hemisphere. All analyses were done separately for the anterior and posterior auditory cortices. Due to the likely violation of the sphericity assumption in repeated measures designs, MANOVAs were employed in order to avoid the inflated Type I error rate associated with the univariate ANOVA when the sphericity assumption is not met.

## 2.6. Designation of anatomical structures

The applied procedure allowed us to derive stereotactic co-ordinates of peak activation for each contrast. However, because we used the Montreal Neurological Institute (MNI) template for normalization, the calculated co-ordinates do not exactly match the original co-ordinates given by Talairach and Tournoux. The reason for that incongruence is the slightly larger MNI template. Therefore, we cross-checked the location of peak activations by referring to the MNI brain. In recognition of the limitations of this technique (both the normalization in general and the usage of the MNI brain in particular), we have taken into account both the primary and the subsidiary contrast peaks as detected through SPM and have illustrated the contiguous

voxels that exceed a Z-statistic of 3.09 for regions showing significant changes in the mean hemodynamic signal. All reported activations are significant at the  $p < 0.001$  level (uncorrected).

## 3. Results

### 3.1. Performance data

Although performance data is missing from one subject and incomplete for another due to technical difficulties, the appropriate degrees of freedom are reported. Percentage of correct responses per condition ranged from 55% in the ‘detect power’ condition to 72% for the ‘detect bower’ condition and 57% in the ‘detect happy’ condition to 75% in the ‘detect sad’ condition. These data were analyzed using a within-subjects, simple repeated measures multivariate analysis of variance (MANOVA). Differences in performance are reflected in a statistically significant difference in the number of correct responses between conditions ( $F(3,6) = 15.4$ ,  $p = 0.003$ ). Post-hoc contrast analyses revealed this difference to be due to significantly better performance during the ‘sad’ condition, compared to both the ‘power’ and ‘happy’ conditions ( $F_s(1,8) > 12.5$ ,  $p_s < 0.009$ ). Data for the performance of both word conditions compared to both emotion conditions (average word vs. average emotion detection) revealed no difference between average performance in the detection of emotion vs. the detection of words,  $t(8) = 0.61$ ,  $p = 0.56$ . There were

Table 1

Regions selectively activated in the contrasts of (A) sad vs. happy, (B) happy vs. sad, (C) sad vs. verbal, (D) verbal vs. sad, (E) happy vs. verbal, (F) verbal vs. happy, (G) emotion vs. verbal, and (H) verbal vs. emotion conditions. Stereotaxic center-of-mass coordinates refer to medial–lateral position (x) relative to midline (positive = right), anterior–posterior position (y) relative to the anterior commissure (positive = anterior), and superior–inferior position (z) relative to the commissural line (positive = superior). R, L: right and left hemisphere. All  $p$ -values are corrected for multiple comparisons. The interpretation of results is based on the voxels thresholded according to the test for spatial extent

Contrast	Anatomical Regions	x	y	z	N of voxel	Size $P(n_{\max} > k)$	Z score	Intensity $P(z_{\max} > u)$
Sad vs. happy	R anterior middle frontal gyrus	36	44	24	23	$p = 0.036$	4.30	$p = 0.005$
Happy vs. sad	No significant voxels							
Bower vs. power	No significant voxels							
Sad vs. verbal	R posterior middle frontal gyrus	40	8	28	24	$p = 0.032$	4.28	$p = 0.004$
	R cingulate gyrus	4	−4	28	25	$p = 0.028$	3.58	$p = 0.004$
Verbal vs. sad	L inferior frontal gyrus	−48	32	−16	31	$p = 0.003$	5.43	$p = 0.002$
	R lingual gyrus/cuneus	8	−76	8	109	$p < 0.001$	4.79	$p < 0.001$
Happy vs. verbal	R superior temporal/inferior parietal lobe	32	−28	28	25	$p = 0.028$	3.94	$p = 0.004$
Verbal vs. happy	L superior temporal gyrus	−60	−48	12	20	$p = 0.053$	4.45	$p = 0.008$
	L precentral gyrus	−60	16	4	21	$p = 0.047$	3.98	$p = 0.007$
Emotion vs. verbal	R inferior frontal gyrus	44	20	16	43	$p = 0.005$	4.32	$p < 0.001$
	L cingulate gyrus	−12	4	20	48	$p = 0.005$	4.04	$p < 0.001$
	R inferior parietal	40	−48	32	33	$p = 0.012$	3.90	$p = 0.001$
Verbal vs. emotion	L inferior frontal gyrus	−48	32	−16	51	$p = 0.002$	5.25	$p < 0.001$
	L middle temporal gyrus	−56	−44	8	63	$p = 0.002$	4.58	$p < 0.001$
	R lingual gyrus/cuneus	8	−76	8	20	$p = 0.054$	4.25	$p = 0.008$

significantly less false alarms during the detection of 'sad' than during the detection of the other three targets ( $F(3,6) = 23.4$ ,  $p = 0.001$ ). Overall, false alarms were very low in all conditions, occurring in less than 4% of all stimulus presentations.

### 3.2. Statistical parametric mapping: activation vs. baseline

The comparison of activation with baseline (ON vs. OFF periods) for all four conditions revealed significant activity primarily in the superior temporal gyrus of both

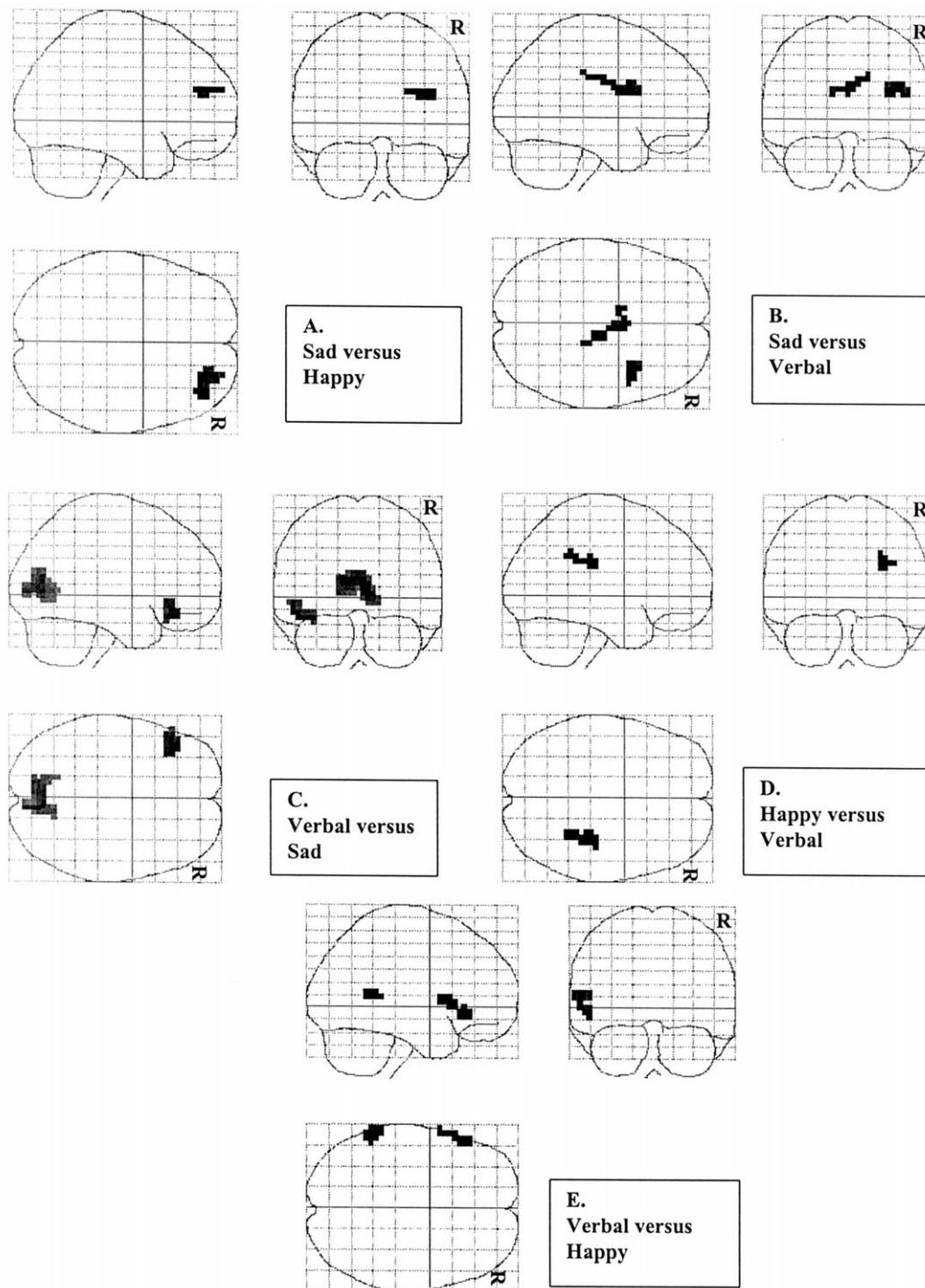


Fig. 2. Schematic representations of mean activated areas on spatially normalised SPM glass brains in comparisons between the following conditions: (A) Sad versus Happy, (B) Sad versus Verbal, (C) Verbal versus sad, (D) Happy versus Verbal, and (E) Verbal versus Happy. Spatial distributions of significantly activated voxels are shown as integrated projections along lateral (top left), coronal (top right) and transverse (bottom) views of the brain (R = right). Pixels were identified as significantly activated if they passed the height threshold of  $Z = 3.09$  and a cluster threshold of  $p < 0.05$ , corrected for multiple comparisons.

hemispheres. These analyses also showed significant bilateral frontal activity primarily within the inferior and middle frontal gyri.

### 3.3. Statistical parametric mapping: contrast analyses

Further analyses focused on differences in activation between the four conditions as well as between the mean activity in the detection of emotion vs. the detection of phonemic stimuli. In the contrast comparing the Sad with the Happy stimulus detection, the right anterior middle

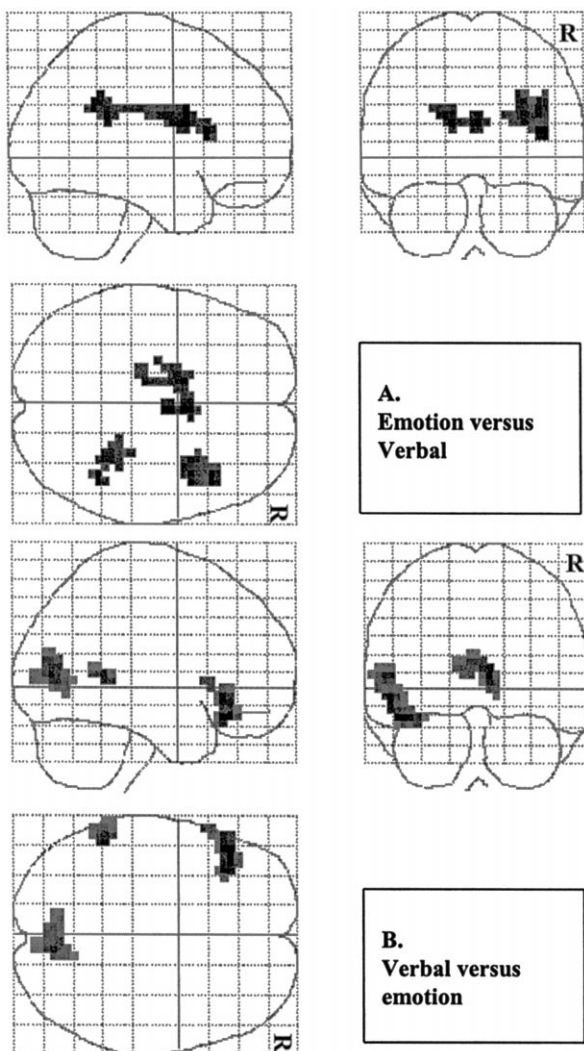


Fig. 3. Schematic representations of mean activated areas on spatially normalised SPM glass brains. Spatial distributions of significantly activated voxels are shown as integrated projections along lateral (top left), coronal (top right) and transverse (bottom) views of the brain (R = right). (A) Emotion–Verbal detection; (B) Verbal–Emotion detection. Pixels were identified as significantly activated if they passed the height threshold of  $Z = 3.09$  and a cluster threshold of  $p < 0.05$ , corrected for multiple comparisons. The maximal foci of activation and the associated  $Z$ -values are given in Table 1.

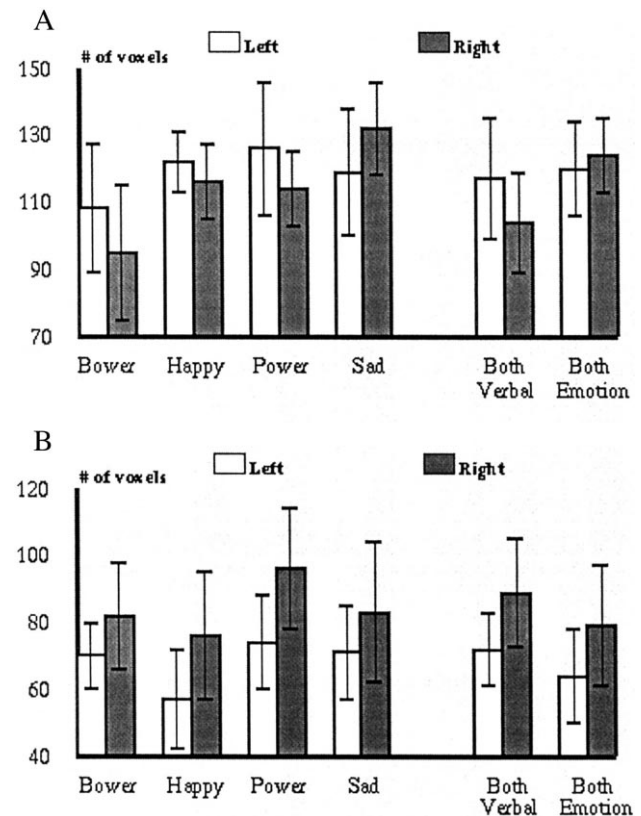


Fig. 4. (A) Mean number of activated voxels in anterior auditory and (B) posterior auditory cortex during all conditions and mean activity during both emotion and both verbal tasks.

frontal gyrus was the only significant area of activation (see Table 1, Fig. 2A). The opposite contrast comparing the Happy with the Sad condition revealed no significant activity. Comparison of the Sad condition with the mean of both verbal conditions (Bower and Power) showed significant activity in the right posterior middle frontal gyrus, as well as in the right cingulate gyrus (Fig. 2B). Fig. 2C shows the contrast comparing both verbal conditions with the Sad condition resulting in significant activation in both the left anterior inferior frontal gyrus as well as the right lingual gyrus and cuneus. The detection of Happy compared with both verbal conditions resulted in significant activity only at the conjunction between the superior temporal lobe and inferior parietal cortex (Fig. 2D). The opposite contrast, shown in Fig. 2E, compares the activity of both verbal conditions with the Happy condition and shows activity in both the left superior temporal gyrus and the left precentral gyrus.

By weighing each condition equally, it was possible to compare both Emotion conditions with both Verbal conditions. The detection of emotion vs. detection of words resulted in significant activity in the right inferior frontal lobe (pars opercularis), the left cingulate gyrus, as well as the right inferior parietal lobe (see Table 1 and Fig. 3).

Comparison of both verbal vs. both emotional conditions revealed significant activation in the left inferior frontal lobe (pars orbitalis), the left middle temporal gyrus, and right extrastriate cortex including the lingual gyrus and cuneus (see Table 1 and Fig. 3).

### 3.4. Anterior auditory cortex activity

Fig. 4A and B show the mean number of activated voxels for the anterior and posterior auditory cortex, respectively, for all conditions as well as the mean activity in both verbal and both emotional detection tasks. There was greater activity in the right anterior auditory cortex during emotion detection compared to verbal detection, as evidenced by a significant task (Emotion or Verbal) by hemisphere interaction,  $F(1,9) = 5.3$ ,  $p < 0.05$ . Post-hoc analysis revealed this difference to be primarily due to significantly greater right hemisphere activity during emotion detection compared to verbal detection conditions,  $t(9) = 1.98$ ,  $p = < 0.05$  (one-tailed). There was not, however, a significant main effect of either task ( $F(1,9) = 2.0$ ,  $p = 0.19$ ) or hemisphere,  $F(1,9) = 0.23$ ,  $p = 0.64$ .

### 3.5. Posterior auditory cortex activity

The analysis of posterior auditory cortex activity resulted in a main effect of hemisphere,  $F(1,9) = 6.1$ ,  $p = 0.036$ , with the right posterior auditory cortex having greater activity in both the emotion and verbal detection tasks. There was not, however, a main effect of task ( $F(1,9) = 1.8$ ,  $p = 0.22$ ), nor a task by hemisphere interaction, ( $F(1,9) = 0.11$ ,  $p = 0.75$ ).

## 4. Discussion

This study sought to identify areas of the intact human brain involved in the detection of emotional prosody using fMRI. The tasks employed involved the perceptual detection of expressed emotion (happy and sad) and phonemes (the initial consonant sounds of 'bower' and 'power') using the same stimuli for all tasks, but with differing task instructions. These subtraction analyses reveal areas likely to be involved in the processing of both emotional intonation and phonemic discrimination. The contrasts comparing language detection with emotion detection resulted in significantly lateralized activity in the frontal lobes, with increased right frontal activity during emotional detection and increased left frontal activity during verbal detection. Additionally, specific analysis of the anterior auditory cortices of both hemispheres revealed greater right hemisphere activation during detection of emotion compared to activity associated with verbal detection. These findings show a close correspondence with both clinical work on

the lateralization of emotion detection following neural insult [see Ref. [47]] and functional neuroimaging studies of emotion detection in normal individuals [20,23].

Bilateral activity was documented during the detection of both emotion and verbal characteristics in language compared to baseline conditions (On vs. Off activity). This activity was especially evident in the temporal and the frontal lobes. The increased activity in the temporal lobes, specifically in the superior temporal gyrus, is consistent with previous research with fMRI in which auditory stimulation results in significant activity in the auditory cortex in spite of the noise produced by the scanner itself [2,26,52]. Additionally, the frontal activity during both tasks is consistent with previous work suggesting a functional connection between the auditory cortex and frontal association cortex (i.e., inferior frontal cortex) involved in the retrieval and rehearsal of auditory information-auditory working memory [10,57–60].

While the current findings are in accordance with the hemispheric asymmetries reported in the clinical literature on aphasia and aprosodia, the localization of activity within the frontal lobes during the tasks is counter to the observations from clinical populations which would predict differences in temporoparietal activity [see Refs. [47,49]]. This discrepancy most likely arises from the inherent differences between experiments involving patients with specific lesions and functional imaging experiments. While studies involving patients with lesions are able to demonstrate those areas *critical* for a specific function, functional neuroimaging studies are only able to demonstrate those areas that are *involved* in a specific function. It must be noted that the frontal activation reported here is based on relative differences in the BOLD response between two language tasks which differ only in the task instructions: to listen for a word or an emotion. Numerous investigators have shown that the left frontal lobe is involved in both phonetic discrimination [3–5,57–60] and the comprehension of language [4].

The lateralized frontal lobe activity resulting from the contrasts of emotional detection with verbal detection is consistent with previous PET studies examining the detection of emotional prosody in spoken sentences [20] and comparing prosodic recognition with recognition of speakers [23]. George and colleagues [20] documented significant activity in the right prefrontal cortex during the perception of emotional prosody in a sentence compared with a control task. In the same study, during the perception of verbal content in spoken sentences, significantly more activity in the left than right frontal cortex was documented. While the task employed in the present investigation involved only emotional detection in words and the task used by George et al. involved emotion detection as well as a verbal response, the congruence between findings from the two studies is noteworthy. The activity documented by Imaizumi et al. [23] is similar to those described in the current investigation. These authors also



documented significant right prefrontal cortex activity during emotional detection in language.

The right frontal cortex activity during the detection of emotion vs. phonemes in the present study is consistent with several functional neuroimaging studies of emotion [30–33]. Morris et al. documented significant right frontal cortex activity while subjects viewed faces with fearful or happy expressions compared to neutral faces [33]. Similarly, Lane et al., [30] report right frontal activation in participants watching emotion-generating film clips compared to a control, neutral film. Interestingly, this right frontal activity was shown while participants watched films associated with the emotions of happiness, sadness, and disgust [30]. Taken together, the current findings and those from these studies of emotion suggest a role of right frontal cortex activity in both the experience and detection of emotion.

Comparison between the detection of a sad vs. a happy emotional tone resulted in significant activity in the right middle frontal gyrus. This finding is consistent with work in the neuropsychology and neurophysiology of emotion [7,9,50]. While the view is highly controversial [see Ref. [48]], the valence hypothesis suggests a lateralization of emotional processing; negative emotions are thought to be processed in the right hemisphere while more positive emotions are processed in the left hemisphere. Specifically, it is postulated that the activity of the anterior frontal cortex of each hemisphere is involved in the processing of the emotions. This view of emotional processing would also predict that the detection of happy versus the detection of sad would result in significant left hemispheric activity. Data from this analysis do not support this prediction, however, as no significant activity was documented in this comparison.

One possible interpretation of these data is that the areas activated during the performance of these tasks are merely those areas involved in different types of auditory detection. Specifically, perhaps the discrimination of emotion could be an example of pitch detection and the detection of the initial consonant in the verbal task is an example of duration detection. This interpretation fits within a model of hemispheric involvement in auditory perception, namely that the left hemisphere is specialized for the analysis of auditory duration while the right hemisphere is more involved in the detection of pitch contours [see Refs. [35,36,46,59]]. The correspondence between the current findings and those of Zatorre et al. [57–60] demonstrating lateralized frontal activity during phonetic and pitch discrimination further support this interpretation. Although the variables of duration and pitch were not specifically manipulated in the current investigation, they are considered to be the primary auditory cues in prosodic expression [see Ref. [51]]. Perhaps the detection of emotional prosody in language is a subset of pitch contour processing and these two phenomena share a common neural basis. Future studies should explicitly manipulate

these auditory characteristics to better understand the relationships among pitch and duration and the detection of prosody in spoken language.

In the comparisons between verbal and emotional conditions, bilateral activity in the lingual gyrus and cuneus was detected (with the focus of greatest activity in the right hemisphere), specifically during the verbal vs. sad and verbal vs. emotion conditions. This activation pattern is unexpected in light of the current literature on the cortical substrates of phonemic processing which would predict activation primarily in the left hemisphere and specifically in the left frontal, temporal, and parietal lobes [4,5]. A recent study showing a similar activation pattern in these extrastriate regions suggests a possible interpretation of the current results. Platel et al. [41], in a study of different aspects of music perception, documented activity in the cuneus during pitch detection. Among several tasks employed by the authors, this pitch detection task was carried out with the least success among the participants. The authors discuss this finding in terms of a cognitive strategy carried out by the participants involving mental imagery. Activation of the extrastriate area is often associated with visual mental imagery [12,29]. Perhaps our results also reflect a visual imagery strategy by the participants specifically during the ‘detect power’ condition. Performance data and debriefing of participants illustrate that the ‘detect power’ condition was the most difficult among all the conditions. We speculate that this pattern of activation may reflect not information about the physical properties of the target stimulus itself, but a cognitive strategy used by the participants in an attempt to improve their performance.

In two comparisons (sad vs. verbal and emotion vs. verbal), bilateral activity in the middle cingulate gyrus was detected (see Table 1 and Figs. 2 and 3). The cingulate gyrus, and specifically the anterior portion, has been the focus of much research in the study of attention [42,43] and emotion [30,31]. Work in the area of attention has suggested that the cingulate gyrus is involved in selection of appropriate responses [43]. Cingulate activity has been documented during the experience of emotion [33], and has been found to be associated with individual differences in emotional experience [31]. The cingulate activity in the current study was documented only during the detection of emotion; in both the sad vs. verbal and emotion vs. verbal contrasts. While this activity was not in the anterior cingulate (the area of the gyrus most associated with emotion), the finding of cingulate activity only during emotion detection suggests that the middle portion of the cingulate gyrus may also be involved in the detection of emotion as well.

Specific analysis of the auditory cortex revealed increased right hemisphere activity during the detection of emotion compared to activity in the same region during phonemic detection. It is important to note that the only difference between these conditions was the instructional set, as both tasks featured the presentation of identical

stimuli. Previous work with neuroimaging of the auditory cortex has shown that activation can be altered by stimulus intensity [26], rate and duration of stimulus presentation [2,27], attention [25,34] and stimulus type [22]. Results from this study extend these findings by illustrating alterations in the BOLD response of the auditory cortex merely by instructions to attend to different aspects of identical stimuli.

The right posterior auditory cortex showed more activation than the left across all stimulus conditions. This increased activity was not due to stimulus conditions, as there were more activated voxels in the right posterior auditory during both verbal and emotion detection tasks. This finding is in line with work from Binder et al. [3,4]. In one study, the presentation of auditory tones resulted in significantly greater posterior auditory cortex activity in the right compared to the left hemisphere [4]. This is in contrast to the activity during a semantic language task within the same experiment. During this task, more anterior auditory cortex activity was documented. These findings support the role of more posterior areas in auditory, as opposed to semantic processing [3–5].

Performance data in this experiment were comparable to those data reported in an earlier study using the same stimuli in a dichotic listening paradigm [7]. The less than 4% false alarm rate across all conditions illustrates that our subjects could well discriminate each specific target. The data from the current experiment do, however, show lower performance across all conditions compared to the Bryden and McRae study [7]. Two possible reasons for this performance deficit are (1) the current study used an inter-stimulus interval of 1 second whereas the previous study allowed a longer interval of 3 s, giving their participants longer to perform the task and (2) potential effects of the experimental setting and ambient scanner noise. Subjects were asked to lie in the scanner for up to a full hour while their heads were held in place to control for movement artifact. Possible effects of experimental setting must be considered when evaluating performance on a complex task such as the one used in the present study. The scanning protocol was designed to circumvent any effect of ambient scanner noise on performance by using a TR time of 6 s and presenting stimuli only when scans were not being made [see Ref. [52] for a more thorough discussion of these issues]. However, one cannot discount the possible carryover effects of the scanner noise on auditory task performance.

The present study illustrates significant lateralization of cortical activity during the perception of both emotional prosody and the perception of verbal characteristics of words. Specifically, this laterality was evident in both the auditory cortex and the frontal lobes. Listening for emotional tone resulted in significantly more activity in the anterior auditory cortex of the right hemisphere compared to the activity of right anterior auditory cortex during the verbal tasks. This work extends previous clinical and

neuroimaging data related to the neuroanatomy of language and emotion processing, but also suggests further work to examine the connections between the temporal and frontal lobes in the processing of language and the continually emerging role of the right hemisphere in language.

## Acknowledgements

The authors would like to thank Dr. Barbara Bulman-Fleming for generously providing the auditory stimuli used in this experiment and Dr. Oliver Zafiris for contributions to the programming of stimuli presentation. This study was financed by the Deutsche Forschungsgemeinschaft (DFG JA 737/5-1). The guest scientists (TWB, KL, LJ) would like to thank the Research Center Jülich for providing the facilities that made this work possible. Tony Buchanan was supported by a grant from the Research Center Jülich as well as an individual National Research Service Award (# 1 F31 MH 11844-01) from the National Institute of Mental Health, US.

## References

- [1] C.J. Bench, K.J. Friston, K.G. Brown, L. Scott, R.S.J. Frackowiak, R.J. Dolan, The anatomy of melancholia. Abnormalities of regional cerebral blood flow in major depression, *Psychol. Med.* 22 (1992) 607–615.
- [2] J.R. Binder, S.M. Rao, T.A. Hammeke, J.A. Frost, P.A. Bandettini, J.S. Hyde, Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex, *Cognit. Brain Res.* 2 (1994) 31–38.
- [3] J.R. Binder, J.A. Frost, T.A. Hammeke, S.M. Rao, R.M. Cox, Function of the left planum temporale in auditory and linguistic processing, *Brain* 119 (1996) 1239–1247.
- [4] J.R. Binder, J.A. Frost, T.A. Hammeke, R.M. Cox, S.M. Rao, T. Prieto, Human brain language areas identified by functional magnetic resonance imaging, *J. Neurosci.* 17 (1997) 353–362.
- [5] J.R. Binder, Neuroanatomy of language processing studied with functional MRI, *Clin. Neurosci.* 4 (1997) 87–94.
- [6] G. Bottini, R. Corcoran, R. Sterzi, E. Paulesu, P. Schenone, P. Scarpa, R.S.J. Frackowiak, C.D. Frith, The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study, *Brain* 117 (1994) 1241–1253.
- [7] M.P. Bryden, L. McRae, Dichotic laterality effects obtained with emotional words, *Neuropsychiatry Neuropsychol. Behav. Neurol.* 1 (1989) 171–176.
- [8] M.B. Bulman-Fleming, M.P. Bryden, Simultaneous verbal and affective laterality effects, *Neuropsychologia* 32 (1994) 787–797.
- [9] R.J. Davidson, Anterior cerebral asymmetry and the nature of emotion, *Brain Cogn.* 20 (1992) 125–151.
- [10] J.F. Demonet, F. Chollet, S. Ramsay, S. Cardebat, J.L. Nespoulous, R. Wise, A. Rascol, R.S.J. Frackowiak, The anatomy of phonological and semantic processing in normal subjects, *Brain* 115 (1992) 1753–1768.
- [11] A.C. Evans, D.L. Collins, S.R. Mills, E.D. Brown, R.L. Kelly, T.M. Peters, 3D statistical neuroanatomical models from 305 MRI volumes, *Proc. IEEE — Nuc. Sci. Symp. Med. Imag.*, 1993, pp. 1813–1817.

- [12] P.C. Fletcher, C.D. Frith, S.C. Baker, T. Shallice, R.S.J. Frackowiak, R.J. Dolan, The mind's eye-Precuneus activation in memory-related imagery, *Neuroimage* 2 (1995) 195–200.
- [13] K.J. Friston, P. Zeigler, R. Turner, Analysis of fMRI time-Series, *Hum. Brain Mapp.* 1 (1994) 153–171.
- [14] K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta, A.C. Evans, Assessing the significance of focal activations using their spatial extent, *Hum. Brain Mapp.* 1 (1994) 210–220.
- [15] K.J. Friston, J. Ashburner, J.B. Poline, C.D. Frith, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, *Hum. Brain Mapp.* 2 (1995) 165–189.
- [16] K.J. Friston, A.P. Holmes, K.J. Worsley, J.B. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, *Hum. Brain Mapp.* 2 (1995) 189–210.
- [17] K.J. Friston, A.P. Holmes, J. Ashburner, J.B. Poline, World Wide Web <http://www.fil.ion.ucl.ac.uk/spm>, 1996.
- [18] F. Galaburda, Sanides, Cytoarchitectonic organization of the human auditory cortex, *J. Comp. Neurol.* 190 (1980) 597–610.
- [19] P.J. Gannon, R.L. Holloway, D.C. Broadfield, A.R. Braun, Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homologue, *Science* 279 (1998) 220–222.
- [20] M.S. George, P.I. Parekh, N. Rosinsky, T.A. Ketter, T.A. Kimbrell, K.M. Heilman, P. Herscovitch, R.M. Post, Understanding emotional prosody activates right hemisphere regions, *Arch. Neurol.* 53 (1996) 665–670.
- [21] W.D. Hopkins, L. Marino, J.K. Rilling, L.A. MacGregor, Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging, *NeuroReport* 9 (1998) 2913–2918.
- [22] K. Hugdahl, K. Bronnick, S. Kyllingsbaek, I. Law, A. Gade, O.B. Paulson, Brain activation during dichotic presentations of consonant-vowel and musical instrument stimuli: a <sup>15</sup>O-PET study, *Neuropsychologia* 37 (1999) 431–440.
- [23] S. Imaizumi, K. Mori, S. Kiritani, R. Kawashima, M. Sugiura, H. Fukuda et al., Vocal identification of speaker and emotion activates different brain regions, *NeuroReport* 8 (1997) 2809–2812.
- [24] L. Jäncke, G. Schlaug, Y. Huang, H. Steinmetz, Asymmetry of the planum parietale, *NeuroReport* 5 (1994) 1161–1163.
- [25] L. Jäncke, S. Posse, N.J. Shah, T. Noesselt, N. Schmitz, H.W. Müller-Gärtner, Attentional factors modify the BOLD response in human auditory cortex to auditory stimuli, *Neuroimage* 5 (1997) 191, abstract.
- [26] L. Jäncke, N.J. Shah, S. Posse, M. Grosse-Ruyken, H.W. Müller-Gärtner, Intensity coding of auditory stimuli: an fMRI study, *Neuropsychologia* 36 (1998) 875–883.
- [27] L. Jäncke, T.W. Buchanan, K. Lutz, K. Specht, S. Mirzazade, N.J. Shah, The time course of the BOLD response in the human auditory cortex to acoustic stimuli of different duration, *Cognit. Brain Res.* 8 (1999) 117–124.
- [28] J.H. Kaas, T.A. Hackett, M.J. Tramo, Auditory processing in primate cerebral cortex, *Curr. Opin. Neurobiol.* 9 (1999) 164–170.
- [29] S.M. Kosslyn, M. Behrmann, M. Jeannerod, The cognitive neuroscience of mental imagery, *Neuropsychologia* 33 (1995) 1335–1344.
- [30] R.D. Lane, E.M. Reiman, M.M. Bradley, P.J. Lang, G.L. Ahern, R.J. Davidson, Neuroanatomical correlates of pleasant and unpleasant emotion, *Neuropsychologia* 35 (1997) 1437–1444.
- [31] R.D. Lane, E.M. Reiman, B. Axelrod, L.S. Yun, A. Holmes, G.E. Schwartz, Neural correlates of levels of emotional awareness: evidence of an interaction between emotion and attention in the anterior cingulate cortex, *J. Cogn. Neurosci.* 10 (1998) 525–535.
- [32] J.S. Morris, C.D. Frith, D.I. Perrett, D. Rowland, A.W. Young, A.J. Calder, R.J. Dolan, A differential neural response in the human amygdala to fearful and happy facial expressions, *Nature* 383 (1996) 812–815.
- [33] J.S. Morris, K.J. Friston, C. Buchel, C.D. Frith, A.W. Young, A.J. Calder, R.J. Dolan, A neuromodulatory role for the human amygdala in processing emotional facial expressions, *Brain* 121 (1998) 47–57.
- [34] D.S. O'Leary, N.C. Andreasen, R.R. Hurtig, R.D. Hichwa, G.L. Watkins, L.L. Boles-Ponto et al., A positron emission tomography study of binaurally and dichotically presented stimuli: effects of level of language and directed attention, *Brain Lang.* 53 (1996) 20–39.
- [35] M.D. Pell, S.R. Baum, Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody, *Brain Lang.* 57 (1997) 195–214.
- [36] M.D. Pell, Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours, *Neuropsychologia* 36 (1998) 701–715.
- [37] M. Peters, Description and validation of a flexible and broadly usable hand preference questionnaire, *Laterality* 3 (1998) 77–96.
- [38] S.E. Petersen, P.T. Fox, M.I. Posner, M. Mintun, M.E. Raichle, Positron emission tomographic studies of the cortical anatomy of single word processing, *Nature* 331 (1988) 585–589.
- [39] S.E. Petersen, P.T. Fox, M.I. Posner, M. Mintun, M.E. Raichle, Positron emission tomographic studies of the processing of single words, *J. Cogn. Neurosci.* 1 (1989) 153–170.
- [40] H. Pihan, E. Altenmüller, H. Ackermann, The cortical processing of perceived emotion: a DC-potential study on affective speech prosody, *Neuroreport* 8 (1997) 623–627.
- [41] H. Platel, C. Price, J.C. Baron, R. Wise, J. Lambert, R.S.J. Frackowiak, B. Lechevalier, F. Eustache, The structural components of music perception: A functional anatomical study, *Brain* 120 (1997) 229–243.
- [42] M.I. Posner, Attention: the mechanisms of consciousness, *Proc. Natl. Acad. Sci. USA* 91 (1994) 7398–7403.
- [43] M.I. Posner, S. Dehaene, Attentional networks, *Trends Neurosci.* 17 (1994) 75–79.
- [44] J.P. Rauschecker, Processing of complex sounds in the auditory cortex of cat, monkey, and man, *Acta Otolaryngol. Suppl. (Stockh.)* 532 (1997) 34–38.
- [45] E.M. Reiman, M.J. Fusselman, P.T. Fox, M.E. Raichle, Neuroanatomical correlates of anticipatory anxiety, *Science* 243 (1989) 1071–1074.
- [46] D.A. Robin, D. Tranel, H. Damasio, Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions, *Brain Lang.* 39 (1990) 539–555.
- [47] E.D. Ross, The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere, *Arch. Neurol.* 38 (1981) 561–569.
- [48] E.D. Ross, R.W. Homan, R. Buck, Differential hemispheric lateralization of primary and social emotions, *NNBN* 7 (1994) 1–19.
- [49] E.D. Ross, R.D. Thompson, J. Yenkosky, Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions, *Brain Lang.* 56 (1997) 27–54.
- [50] H.A. Sackheim, M.S. Greenberg, A.L. Weiman, R.C. Gur, J.P. Hungerbühler, N. Geschwind, Hemispheric asymmetry in the expression of positive and negative emotions: neurologic evidence, *Arch. Neurol.* 39 (1982) 210–218.
- [51] K.R. Scherer, Vocal affect expression: a review and a model for future research, *Psychiatr. Bull.* 99 (1986) 143–165.
- [52] N.J. Shah, L. Jäncke, M. Grosse-Ruyken, H.W. Müller-Gärtner, Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination, *J. Magn. Reson. Imaging* 9 (1999) 19–25.
- [53] H. Steinmetz, J. Rademacher, L. Jäncke, Y. Huang, A. Thron, K. Zilles, Total surface of temporoparietal intrasylvian cortex: diverging left-right asymmetries, *Brain Lang.* 39 (1990) 357–372.
- [54] H. Steinmetz, Structure, function and cerebral asymmetry: in vivo morphometry of the planum temporale, *Neurosci. Biobehav. Rev.* 20 (1996) 587–591.
- [55] D. Stiller, B. Gaschler-Markefski, F. Baumgart, F. Schindler, C. Tempelmann, H. Heinze, H. Scheich, Lateralized processing of speech prosodies in the temporal cortex: a 3-T functional magnetic resonance imaging study, *MAGMA* 5 (1997) 275–284.
- [56] J. Talairach, P. Tournoux, Co-Planar Stereotaxic Atlas of the Human

- Brain, 3-Dimensional Proportional System: An Approach to Cerebral Imaging, Thieme Medical Publishers, New York, 1988.
- [57] R.J. Zatorre, Pitch perception of complex tones and human temporal-lobe function, *J. Acoust. Soc. Am.* 84 (1988) 566–572.
- [58] R.J. Zatorre, S. Sampson, Role of the right temporal neocortex in retention of pitch in auditory short-term memory, *Brain* 114 (1991) 2403–2417.
- [59] R.J. Zatorre, A.C. Evans, E. Meyer, A. Gjedde, Lateralization of phonetic and pitch discrimination in speech processing, *Science* 256 (1992) 846–849.
- [60] R.J. Zatorre, A.R. Halpern, D.W. Perry, E. Meyer, A.C. Evans, Hearing in the mind's ear: a PET investigation of musical imagery and perception, *J. Cogn. Neurosci.* 8 (1996) 29–46.