Cerebral Processing of Timbre and Loudness: fMRI Evidence for a Contribution of Broca's Area to Basic Auditory Discrimination

Susanne Reiterer · Michael Erb · Wolfgang Grodd · Dirk Wildgruber

Abstract Sound timbre and sound volume processing are basic auditory discrimination processes relevant for human language abilities. Regarding lateralization effects, the prevailing hypotheses ascribe timbre processing to the right hemisphere (RH). Recent experiments also point to a role of the RH for volume discrimination. We investigated the relevance of the RH for timbre and volume processing, aiming at finding possible differences in cerebral representation of these acoustic parameters. Seventeen healthy subjects performed two auditory discrimination tasks on tone pairs, differing either in timbre or volume. FMRI was performed using an EPI-sequence on a 1.5 T scanner. Hemodynamic responses emerged in both tasks within a bilateral network of areas, including cingulate and cerebellum, peaking in primary and secondary auditory cortices (core and belt areas). Laterality analyses revealed a significant leftward dominance at the temporal cortex. Task comparison revealed significant activation within Broca's area during the timbre task and a trend for an increase of right parietal responses during volume processing. These results contribute to a more differentiated view of timbre processing. Additionally to the engagement of the right temporal cortex during processing of musical timbre, there seem to be language related aspects of timbre that are preferentially processed in the left hemisphere. These findings are discussed within the framework of a model of timbre perception comprising two differentially lateralized subprocesses. Processing of spectral cues (harmonic structure) linked to the right hemisphere and processing of temporal cues (i.e. attack-decay dynamics) linked to the left hemisphere. Moreover, activation of Broca's area linked to the timbre task indicates a participation of this area in discriminating phonetic changes of the vowel-like non-speech signals, encouraging the argument that basic acoustic cue processing at a pre- or non-speech level is represented within this "classical language area."

Keywords fMRI · Timbre · Volume · Basic auditory processing · Language processing · Subcortical · Cerebellum · Broca's area · Sound perception · Auditory cortex

S. Reiterer (✉) · M. Erb · W. Grodd
Department of Neuroradiology, Section of Experimental MR of the CNS, University of Tübingen,
Hoppe Seyler Str. 3,
72076 Tübingen, Germany
e-mail: Susanne.Reiterer@med.uni-tuebingen.de

S. Reiterer
Institute for Natural Language Processing, Experimental Phonetics Group, University of Stuttgart,
Azenbergstr. 12,
70174 Stuttgart, Germany

D. Wildgruber
Department of Psychiatry, University of Tübingen,
Osianderstr. 24,
72076 Tübingen, Germany

Introduction

Amongst the most salient basic acoustic parameters, like sound duration and pitch, there are parameters which have been experimentally underrepresented within neuropsychological and neuroimaging research, but which are at the same time highly relevant for human cognitive phenomena, such as language and music perception: namely timbre (a subset of sound quality) and volume (sound intensity, loudness). Sounds may be generally characterized by duration, pitch, loudness and quality. Sound "quality" more generally, or "timbre" more specifically, describes those

characteristics which allow the ear to distinguish sounds which have the same pitch and loudness (Grey 1977). Regarding timbre, the neurophysiological or psychological studies published concentrate mostly on timbre as a discriminating feature for the perception of music, thus investigating mainly aspects of "musical timbre," with experiments on the discrimination of musical instruments or melodies differing in timbre. Results preponderantly show right hemisphere (RH) involvement for musical timbre (Boucher and Bryden 1997; Halpern et al. 2004; Samson and Zatorre 1994; Platel et al. 1997).

Timbre is mainly determined by the harmonic content of a sound and the dynamic characteristics of the sound such as vibrato and the attack-decay envelope of the sound. Especially for sustained tones, the most important of these factors is harmonic content, the number and relative intensity of the upper harmonics present in the sound. The right hemisphere has been observed to be specifically sensitive for processing of these spectral sound features (Menon et al. 2002; Johnsrude et al. 1997; Zatorre et al. 2002; Jäncke et al. 2002; Warren et al. 2005).

Most of these studies used stimulus sounds in strings of longer events, mostly melodies, where sounds of different instruments had to be discerned. Studies using isolated tones for timbre differentiation, however, brought more divergent results in terms of differential lateralization. Applying a dichotic listening paradigm, Brancucci (Brancucci and San Martini 2003) found significant right hemisphere activation in response to timbre differences produced by dissimilar amplitude envelopes of complex tones (timbre fluctuations of a steady state complex tone), and Dehaene-Lambertz (2000) in an event-related potentials (ERP) study with infants found preferential left hemisphere (LH) engagement underlying perception of tones changing in number of harmonics (timbre change). Taking into account these findings, we tried to re-examine the laterality of timbre processing outside of a music context, hypothesizing more rightward activations for sound timbre processing.

Considering cerebral representation of sound intensity/ volume/loudness processing, the few studies investigating it so far showed in the majority a right hemisphere involvement in volume processing. In a study by Belin et al. (1998), a right hemisphere fronto-parietal network was shown to be involved in sound intensity discrimination. Other works (Lasota et al. 2003; Opitz et al. 2002; Mustovic et al. 2003) revealed bilateral auditory cortex (AC) areas, such as supratemporal gyrus (STG) and Heschl's gyrus (HG) in addition to right hemispheric areas, like the temporo-parietal junction area in response to loudness and silence (Mustovic et al. 2003) and RH inferior frontal parts in response to intensity change detections (Opitz et al. 2002). Using dichotic listening to detect hemispheric asymmetries, Brancucci et al. (2005) employed complex synthesized tones as well as natural voice speech syllables of varying input volume and found a right hemisphere asymmetry for both stimulus types, speech and non-speech. Since no strong left hemisphere advantages for volume processing have been reported in the literature, we hypothesized that volume processing in our sound discrimination experiment would provoke either right hemispheric or bilateral involvement.

In summary, neural correlates of volume and timbre processing are still not well understood, and lateralization effects are discussed controversially. A major confounding factor is that stimuli and tasks (either comprising longer or shorter stimuli, trains of sounds, single tones, embedded into a musical, speech or basic acoustic context, manipulations at spectral, temporal or other levels etc.) vary enormously across different studies. To circumvent this problem we looked at differences in timbre as well as volume discrimination within one and the same experimental paradigm. We aimed at comparing timbre and volume processing to detect possible differences between the two categories within and between the hemispheres (laterality phenomena) by employing the same task: a paired sound discrimination paradigm. We applied 200 ms difficulty— varied harmonic content—or volume-manipulated synthesized acoustic signals and were interested in (1) to what extent the two basic acoustic processes differ in terms of their hemispheric lateralization biases and (2) to what extent the task difficulty had an impact on the lateralization and representation of the two parameters.

Methods

Subjects

In this study 17 healthy volunteers having given written informed consent were investigated [8 male, 9 female, mean age 25.2 years, range 18–31 years]. The study protocol was approved by the local ethics committee as meeting the requirements for the Code of Ethics according to the declaration of Helsinki for investigations on human subjects. All subjects were strongly right handed as assessed by the Edinburgh handedness scale (Oldfield 1971; laterality index >90%). All participants had comparable educational status and no psychiatric, neurological or hearing disorders.

Task

The experimental paradigm consisted of two forced choice paired-discrimination tasks: difficulty-varied timbre (sound quality) and difficulty-varied volume (sound intensity) discrimination. Subjects had to either discern the "brighter"

(timbre task) or the "louder" (volume task) of the two stimuli, binaurally presented with a fixed delay of 500 ms.

In several separate behavioural pilot experiments on different subjects ($n=30$) outside the scanner, the stimulus pairs were pre-tested for their discriminability, difficulty-matched according to the resulting performance scores and then used in the fMRI experiment.

Stimuli

The stimulus material comprised synthesized four-component complexes where the components were formant-like in spacing, comprising the four formant components: F1–F4: 500, 1500, 2500, 3500 Hz (for the reference tone). All stimuli were of 200 ms duration with a pitch being superimposed on the sound by modulating the amplitude across all components with a periodicity of 150 Hz (F0, close to typical female fundamental in speech). Stimuli were synthesized using a vowel synthesizer based on formant sinusoids (Hertrich and Ackermann 1999) and sounded like typical computer-generated sine-wave speech with additional pitch and changes in quality and volume. Stimuli were manipulated either in timbre or volume and were all within the language-specific spectral range. The timbre variations were induced by variation of the gaps between the 4 formant-like components (formant frequencies). The first component was always set at 500 Hz, and the gaps between successive components (F1–F2, F2–F3, F3–F4) were varied in 29 steps ranging from 500 to 1500 Hz per formant gap. Variation of sound volume was also carried out in 29 different steps, ranging from 8,000 to 32,000 arbitrary loudness units, each signal being compared to the reference tone with 16,000 arb. loud. units (Fig. 1). During one session of the experiment 57 timbre-varied pairs of varying difficulty were presented. During the other session 57 volume-varied pairs were presented. Order of sessions and stimuli pairs were balanced and pseudor-andomized across subjects. Each manipulated tone was

compared to the reference tone described above (see also Fig. 1). To avoid habituation or "priming" effects of saliency or markedness, we randomised the order of the reference tone. In half of the cases the reference tone came first, and in the other half it came as the second tone.

fMRI recording

For the acoustic stimulation special MR-compatible ear-phones, based on piezo-electric signal transmission, were used (Jäncke et al. 2002). After task instruction, earphone volume was adjusted to the individual subjects' needs during a test scan. We had two conditions which each lasted 8 min and we recorded 256 volumes within each session. In each of the two conditions 57 differing stimulus pairs were presented with temporal jittering, randomized according to their physical differences (delta timbre or delta volume, i.e. the difficulty level). To avoid systematic acoustic interference between scanner noise and test stimuli, intervals between stimuli were varied between 8.25 and 23.25 s hereby introducing a temporal jittering.

Subjects pressed a button (left and right hand equally distributed) to indicate their decision after presentation of each stimulus pair. Event-related fMRI was performed using an EPI sequence (1.5 Tesla, Siemens Vision, TR= 3 s, TE=40 ms, FOV=192 mm, 28 axial slices, slice thickness=4 mm, sequential descending order of acquisition, voxel size=3×3 mm, matrix 64×64, flip angle 90°).

Statistical analysis

Preprocessing of functional images included motion correction, slice time correction to the middle slice, normalization into MNI space (Montreal Neurological Institute), and spatial smoothing with a conventionally used standard Gaussian Filter of 10 mm full width at half maximum (FWHM). Four subsequent statistical analyses (random effect analysis) were
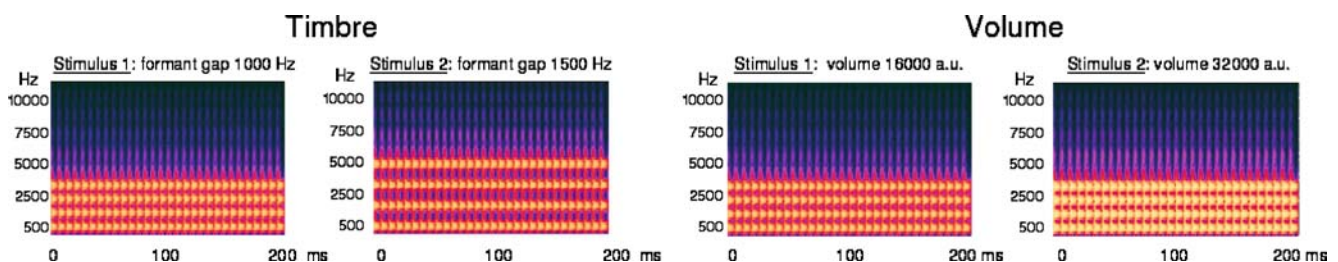

Fig. 1 Synthesis of acoustic stimuli: complex acoustic signals comprising four harmonics (reference tone=500, 1,500, 2,500 and 3,500 Hz) were varied in their timbre or volume. Timbre variation was realized by modulation of formant gaps (500–1,500 Hz). As compared to the reference stimulus (formant gap 1,000 Hz), the difference of formant gaps (delta timbre) between each stimulus and the reference tone thus varied from 0–500 Hz. Loudness of the stimuli was varied from 8,000 to 32,000 arbitrary loudness units (a.u.). The difference of loudness (delta volume) as compared to the reference tone (16,000 a.u.), thus ranged between 0 and 16,000 a.u.. Fifty-seven stimulus pairs, exclusively differing in one of these parameters, were used for the respective discrimination task. *Left*: pair of timbre-varied stimuli (formant gap: 1000 vs.1500 Hz; delta timbre=500 Hz); *right*: pair of volume-varied sample stimuli (16,000 vs. 32,000 a.u; delta volume=16,000 a.u.)

Fig. 2 Behavioural data for discrimination of tone pairs differing in timbre (*left*) or volume (*right*). Hit rates (*y*-axis) are plotted with respect to graded differences of timbre and volume between the two acoustic signals (delta timbre and delta volume, represented on the *x*-axis)
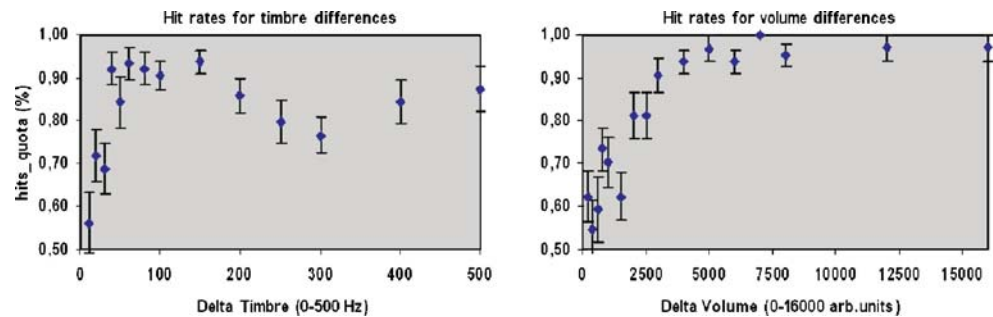
carried out using SPM2 (Wellcome Department of Imaging Neuroscience, London, UK): (1) Analysis of main effects (Comparison of task versus baseline; the used baseline derives from the idling periods (rest) between explicit tasks); (2) Analyses of laterality effects (left versus right hemisphere) to elucidate hemispheric differences (by flipping of the contrast images at the *y*-axis along the direction of the *x*-axis (horizontal axis) and comparing the inverted against the non-inverted images on a voxel-by-voxel basis); (3) Analyses of parametric effects to extract possible task difficulty-related effects (linear correlation between the hemodynamic blood oxygen level dependent (BOLD) response and the behavioural performance scores); (4) Analysis of categorical effects (task versus task comparisons), e.g. timbre versus loudness discrimination.

As the standard criterion of statistical significance, a height threshold at voxel level of $p<0.001$ ($T>3.69$), corrected at cluster level for multiple comparisons $p<0.05$ (extent threshold $k>90$ voxels), was applied.

To increase the sensitivity of statistical analysis we introduced a second cut-off criterion (trend-level), using a height threshold of $p<0.01$ ($T>2.58$) at voxel level, and an additional extension threshold ($k>90$), reaching an uncorrected $p<0.01$ at cluster level.

Finally, the anatomical labelling of the activation maps was performed using Automated Anatomical Labelling (Tzourio-Mazoyer et al. 2002) and Cytoarchitectonic Probability Maps (Morosan et al. 2001; both implemented in the toolboxes available for SPM2.

Results

The behavioural data (Fig. 2) showed hit scores of about 75% for both tasks. The hit scores obtained during the fMRI experiment were within the same range as the performance scores of the preceding behavioural experiments outside the scanner, precluding strong influence of scanner noise on acoustic discriminability. For both tasks, the rate of correctly identified stimuli increased approximately with rising physical difference between the two acoustic signals (delta timbre/volume). Hit rates for timbre

discrimination showed slightly more variation, not reaching the 100% hit quota in the stimuli with larger differences (delta timbre between 200 and 500 Hz).

Main effects

The hemodynamic responses for the main effects (task versus baseline) in both tasks were largely similar and emerged within a widespread bilateral network of cortical and subcortical areas, including mainly temporo-parietal and frontal areas, as well as thalamus, basal ganglia, cingulate and cerebellum, peaking bilaterally in primary and secondary auditory cortices (core and belt areas) and inferior parietal areas, especially in the volume task. Laterality analyses of both main effects (timbre and volume) showed a circumscribed significant left temporal cluster: left posterior supratemporal gyrus and middle temporal gyrus (MTG), covering Heschl's Gyrus, temporal plane and BA 41 within the primary acoustic cortex and portions of the insula (see Fig. 3, lower rows of left and middle panel, Table 1). Thus, similar main effects resulted also in a similar leftward laterality effect for main BOLD effect in sound timbre as well as volume discrimination.

Parametric effects

As for the linear correlations between the BOLD activity and the performance measures (hit rates), the parametric analyses only showed trends for activations within the right inferior temporal gyrus (ITG) as well as parts of the left, but mainly right cerebellum (lobule 6 and 8) with increasing success in timbre discrimination (Fig. 4 upper row in left-hand panel "Parametric Effect Quality"). For increasing success in the volume task a trend for a linear relationship to a cluster within the left lentiform nucleus and the posterior parts of right and left cingulum was detected. Laterality analyses upon these parametric effects (see Fig. 4 lower row in left and middle panel) showed a right lateralized cerebellar cluster (in lobule 6) for the parametric effect in successful timbre discrimination and a left lateralized cluster within the area of the left hippocampus for the linear increase with better volume discrimination.
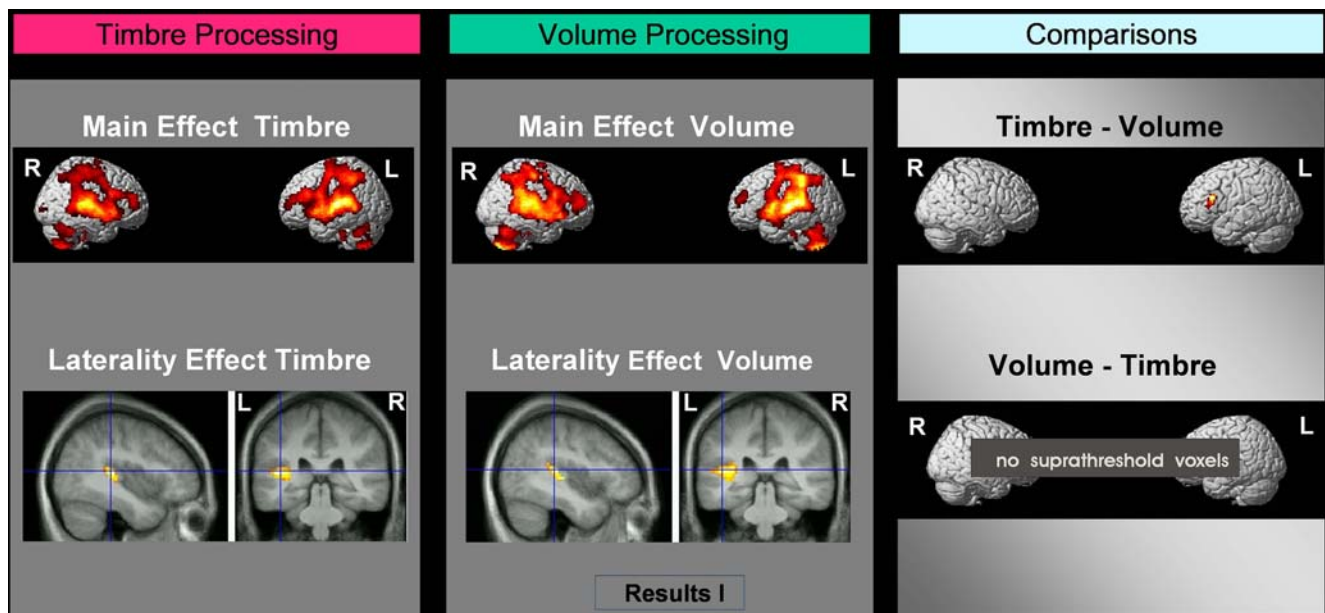
Fig. 3 *Left & middle panel, upper row*: Main effect: cerebral activation during timbre and volume discrimination. Significantly activated areas as compared to rest are projected upon the cortical surface of a template brain. Both tasks yielded a fronto-temporo-parietal activation pattern, with the peak activation residing in left auditory areas. *Left & middle panel, lower row*: Lateralization effect: Voxelwise comparison of the hemispheres against one another, revealing a similar left-temporal lateralization effect for both tasks. *Right-hand panel*: Differential effects (task versus task comparisons) showing a significant left inferior frontal (IFG) activation for timbre>volume and no significant effects for volume > timbre (height threshold: $p<0.001$, corrected at cluster level at $p<0.05$)

Categorical effects

However, the differential effects, i.e. task versus task comparisons of timbre versus volume processing (see Figs. 3 and 4, right panel, Table 2), showed a significant cluster in Broca's area (BA 45, pars triangularis of the left inferior frontal gyrus (IFG), Tzourio-Mazoyer et al. 2002) during timbre processing (timbre>volume) and no activation cluster during

Table 1 This table shows the exact voxel coordinates (according to the convention of the Montreal Neurologic Institute, "MNI") with their respective statistical t-values of the peak activations for: (1) main effects, (2) laterality effects of main effects, (3) parametric effects and (4) laterality effects of parametric effects for each of the two conditions

Timbre processing					Volume processing				
Peak areas	x	y	z	cluster size / peak t	Peak areas	x	y	z	cluster size / peak t
1. Main effects: standard level of significance: (height threshold $p<0.001$, $T>3.69$, corr. at cluster level $p<0.05$)									
L-STG, sup. temp. gyrus	−48	−15	0	1280 / 17.46	R-SMA, suppl. motor area	6	−9	69	464 / 15.18
R-STG, sup. temp. gyrus	51	−21	6	883 / 17.24	L-STG, sup. temp. gyrus	−54	−21	0	683 / 14.66
L-IPG, inf. parietal gyrus	−51	−45	45	81 / 10.40	L-Thalamus	−12	−18	0	341 / 13.46
2. Laterality Effects of Main effects: standard level of significance, as in 1. Main Effects									
L-STG, sup. temp. gyrus	−42	−33	12	107 / 7.28	L-STG, sup. temp. gyrus	−45	−33	12	145 / 6.84
L-Insula	−42	−3	−3	38 / 6.25	L-Heschl's Gyrus	−36	−33	12	33 / 6.11
L-PCG, precentral gyrus	−51	0	30	44 / 5.87	L-PCG, precentral gyrus	−45	0	27	36 / 5.83
3. Parametric effects: lowered level of significance: (height threshold $p<0.01$, $T>2.58$, uncorr. at cluster level $p<0.01$)									
R-ITG, inf. temporal gyrus	45	−12	−27	513 / 4.73	L-Putamen	−27	−6	−6	152 / 5.10
R-Cerebellum, lobule 6	27	−54	−30	119 / 4.26	R-post. Cingulum	3	−57	30	222* / 4.13
L-Cerebellum, lobule 8	−6	−63	−33	123 / 4.03	L-post. Cingulum	−6	−51	21	(222)* / 4.10
4. Laterality Effects of Parametric effects: lowered level of significance, as in 3. Parametric Effects									
R-Cerebellum, lobule 6	24	−54	−30	119 / 4.89	L-Hippocampus	−24	−6	−27	91 / 4.90
	30	−42	−27	(119) / 3.78					

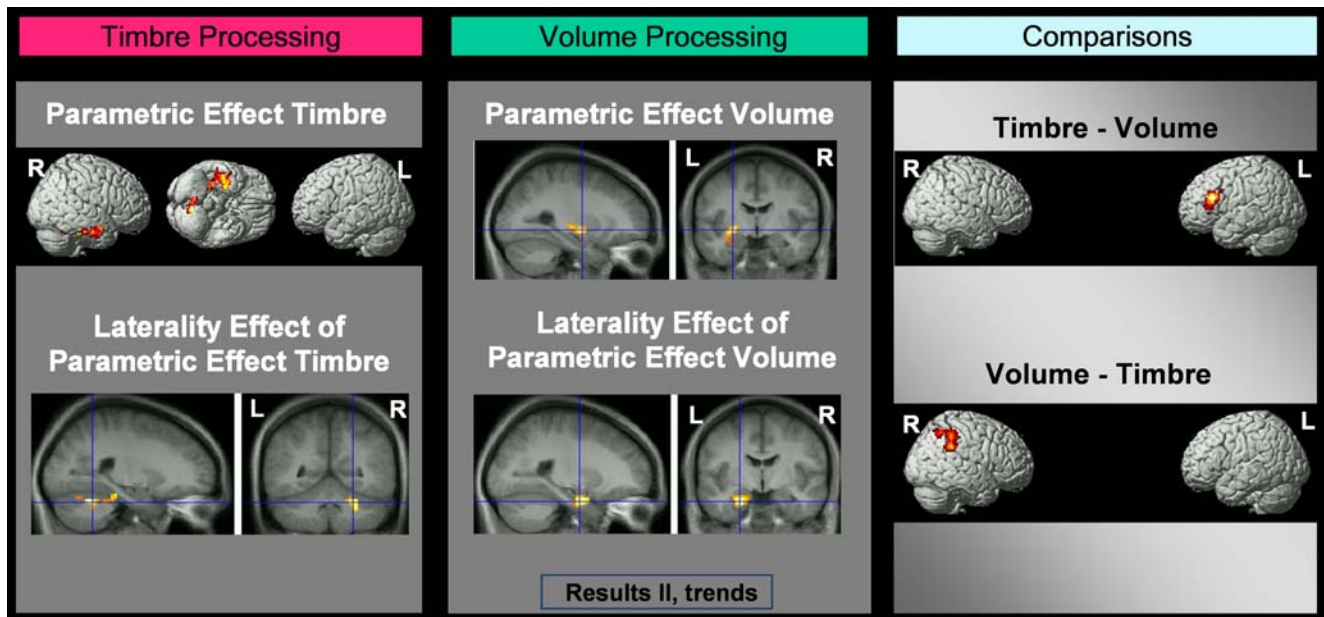Cluster side is given in voxels. *here: cluster size of L cingulum and R cingulum together is 222 voxels.

Fig. 4 This figure represents "trends" of activations, which emerged after applying the more sensitive threshold of $p<0.01$, uncorrected at cluster level (height threshold at voxel level: $p<0.01$ and cluster extension threshold ($k>90$), reaching $p<0.01$ uncorrected at cluster level. *Left & middle panel, upper row*: Parametric Effects: cerebral activation correlated to increase of the behavioral performance score. Accuracy of timbre discrimination was associated with hemodynamic responses within the right inferior temporal gyrus (ITG) as well as parts of the right (and left) cerebellum. Corresponding areas for the volume task were found within the left lentiform nucleus as well as the posterior part of right and left cingulum (not depicted in this brain slice). *Left & middle panel, lower row*: Lateralization effect of parametric effects: activation patterns after comparison of the hemispheres against one another. Laterality analysis of the aforementioned parametric effect showing a right cerebellar cluster during timbre discrimination, and a left hippocampal (parahippocampal) area for the volume task. *Right panel*: Differential effects at a more sensitive level showing a more extended IFG cluster (more extended into the middle frontal area) for timbre>volume and a right inferior parietal cluster (including the right supramarginal gyrus, BA 40, and angular gyrus) for volume > timbre

volume discrimination at the standard level of significance (height threshold $p<0.001$ ($T>3.69$), corrected at cluster level $p<0.05$). However, applying the more sensitive criterion (height threshold at voxel level $p<0.01$ ($T>2.58$) and cluster extension threshold $k>90$, reaching an uncorrected $p<0.01$ at cluster level) revealed a trend for a right inferior parietal cluster, around (above and including) the right supramarginal

gyrus (BA 40; max: $x=60$, $y=-42$, $z=33$) as well as the right angular gyrus for the volume discrimination task (volume>-timbre). In the case of the comparison timbre>volume, the more sensitive analysis (Fig. 4) showed the same results as the conservative one (Fig. 3): an activation in the left inferior frontal gyrus (mainly pars triangularis, BA 45), this time only more extended into the area of the middle frontal gyrus.

Table 2 This table shows the exact voxel coordinates (according to the convention of the Montreal Neurologic Institute, "MNI") with their respective statistical *t*-values of the peak activations for the: Differential (Categorical) Effects (task versus task comparisons) at the (1) standard level of significance as well as (2) the lowered level of significance for each of the two comparisons

Task versus task comparisons:

Timbre—Volume					Volume—Timbre				
Peak areas	x	y	z	cluster size / peak *t* value	Peak areas	x	y	z	cluster size / peak *t* value
1. Categorical Effects at standard level of significance: (height threshold $p<0.001$, $T>3.69$, corr. at cluster level $p<0.05$)									
L-IFG, inf. frontal gyrus	−42	24	21	94 / 6.05	Non significant	–	–	–	–
2. Categorical effects at lowered level of significance: (height threshold $p<0.01$, $T>2.58$, uncorr. at cluster level $p<0.01$)									
L-IFG, inf. frontal gyrus	−42	24	21	268* / 6.05	R-IPG, inf. parietal gyrus	57	−42	45	59* / 6.16
					R—supramarginal gyrus	60	−42	33	59* / 5.56
L-MFG middle frontal gyrus	−42	12	33	268* / 2.75	R—angular gyrus	48	−60	51	19 / 3.78

Cluster size is given in voxels. *here: cluster size of L-IFG and L-MFG together: 268 voxels (59 respectively).

Discussion

Behavioural effects

The behavioural data showed comparable hit scores of about 75% for both tasks. Hit rates for volume discrimination continuously increased with higher differences, whereas timbre discrimination showed slightly more variation (local minimum around 250 Hz distances). The reason for this presumably lies in the phenomenon that formant structure of one sound interacts with the harmonic structure of the other sound (i.e. if common integer multiples appear in the harmonic structure of the other sound), so that common partial tones or common harmonics come into existence (Grey and Gordon 1978), that interfere with discriminability of the perceived timbre. This is also the reason why, for example, the interval of a second on the piano is easier to discriminate than the interval of an octave although the formant structures of the tones in the octave are more distant.

fMRI main effects (task versus baseline)

The results of the main effects show generally similar activations with a left-sided temporal peak cluster within STG/HG/PT for both sound timbre and sound volume processing. By analysing the laterality of this effect, we found generally similar activations for both sound timbre and sound volume discrimination, with a significant hemispheric asymmetry towards a left-sided temporal cluster within STG/HG, on the basis of the detailed voxel per voxel laterality analyses. This left hemisphere bias for both tasks is in contrast with the hypotheses of predominantly RH involvement in timbre processing as well as in volume processing (Samson 2003; Halpern et al. 2004; Belin et al. 1998; Brancucci and San Martini 2003). A left AC activation in Heschl's gyrus in response to increased sound intensity discrimination has so far been reported only as a nonsignificant trend by Lasota et al. (2003). In another experiment by our group (Reiterer et al. 2005) using a similar experimental setup investigating the discrimination of the acoustic parameters pitch and duration, we also observed an unexpected LH asymmetry (left STG/HG) for pitch as well as duration processing when analyzing the main effects (i.e. when task was compared to baseline).

However, a hemispheric bias for the left auditory cortex for the processing of timbre has already been reported in studies (Deike et al. 2004; Menon et al. 2002; Dehaene-Lambertz 2000) also using "longer," sustained stimuli of comparable length (comparable to our stimuli) where mainly the harmonic content was manipulated.

Since we used an active listening paradigm with directed attentional resources towards the acoustic differentiation task, it could be argued that the active attention itself would cause a left lateralization regardless of the acoustic parameter, or even regardless of whether the input signal is speech or non-speech, as for example shown in Hertrich (Hertrich et al. 2003). Upon this point it has to be mentioned that the above cited investigations (Menon et al. 2002; Dehaene-Lambertz 2000) used already a passive listening paradigm and still found this LH engagement. Furthermore, specifically on this question, a recent study (Vihla and Salmelin 2003), comparing cortical processing of attended and non-attended vowels and complex tones, could show that responses were similar during active as well as passive listening. Thus, we tentatively conclude that, the left-bias seems not to have been introduced by attentional constraints. Additionally, we would like to rule out the assumption, that left lateralization could have occurred due to task difficulty. Since no significant correlations of hemodynamic activity within the left temporal regions and difficulty of processing was found, this parameter does not seem to bring about the observed leftward bias.

However, a more likely interpretation of our results seems to be that acoustic parameters with a temporal fine structure which require rapid temporal processing (small time changes) are predominantly processed within the left hemisphere. As already stated above, timbre perception is based on temporal and spectral cues. It is inevitable that both features are always present to some degree in a timbre stimulus, but we think that the stimulus design and task discrimination might have led the subjects to exploit more the temporal than the spectral cues as main behavioural strategy in their decision making in our study. This could explain the observed leftward lateralization including the involvement of Broca's area in the timbre processing. A corresponding result was already reported by Platel et al. (1997) where a rhythm task activated left inferior Broca's area, with extension into the neighbouring insula, suggesting a role for this cerebral region in the processing of sequential sounds.

fMRI categorical effects (timbre vs. intensity discrimination)

Outside of the auditory areas, within the left inferior frontal gyrus, pars triangularis (BA 45, part of Broca's Area) we detected significant activation for the processing of timbre as compared to volume discrimination. This could be due to higher order phenomena as, for example, differences in perception of categorically shifting vowel-like stimuli as opposed to more continuous changes in intensity. In line with this assumption, prior PET and fMRI studies have been linking phonological vowel discrimination to Broca's area (Fiez et al. 1995; Hsieh et al. 2001; Gandour et al.

2002). More specifically, this inferior frontal activity could have resulted on the one hand from a participation of this area in discriminating the "language-related" aspects of the timbre of our vowel-like but at the same time—strictly speaking—non-speech signals, encouraging perhaps the mirror neuron system (Iacoboni et al. 1999, 2005) in the area around Broca to engage in an internal imitation or subarticulation process of the different vowel qualities to achieve better discrimination of the two vowel-like sounds (as the task was to discriminate between the "brighter" and the "darker" of two sounds, which sounded like derivatives of the German "umlaut" vowels /ö/ and /ä/). On the other hand, short temporal discrimination of timbre is also a necessary prerequisite for the perception of vowels and could thus point to phonological processing in left IFG. The role of Broca's area in speech perception and its overlap of function in the form of a "production and perception" network, is by now a well-established view (compare Wilson et al. 2004; Heim et al. 2003; Scott and Johnsrude 2003). Broca's area has also been connected to phonological segmentation processes (Burton et al. 2000) as well as to rapid non-speech frequency changes as exemplified by the use of tonal frequency glides with formant changes (Müller et al. 2001). Broca's area, thus, seems to be involved in basic acoustic timbre discrimination that might be crucial for phonological processing of speech sounds. Although phenomenologically more often attributed to the domain of music, the acoustic property of sound that allows a person to distinguish two sounds when pitch, loudness, and perceived duration remain identical, also allows one to differentiate human voices (during singing and speaking) as well as linguistic phonetic categories, such as for example vowel categories. In the domain of language perception, humans show an impressive ability to both discriminate between and generalize over human speech sounds, by using formants as the critical discriminative cue (Hauser et al. 2002). Thus, we would like to refer to changes in the quality of sound with relevance to language processing as "language-relevant or language-related timbre". Some brain imaging studies have investigated correlates of vowel processing (different in "language-related timbre"), investigating different vowel categories (e.g. /a/, /i/ and /u/). Here MEG source localization (Shestakova et al. 2002; Vihla and Salmelin 2003) resulted in left hemisphere activations. All in all, the role of Broca's area for phonological perception and coding has been consolidated and described by now in various studies (Joanisse and Gati 2003; Huang et al. 2002; Platel et al. 1997). Furthermore, the engagement of Broca's area is well documented and reported almost equally often outside the domain of language, like in music perception (Koelsch et al. 2002; Levitin and Menon 2003) and in motor imitation, action recognition and social intention (Iacoboni et al. 2005).

Summarizing, as we can see from the results of our study and from these various and diverse above cited neuroimaging studies, Broca's area seems to serve a complex heterogeneity of function and all these studies activating Broca's area possibly share one or more similar and specific aspects of stimulus features, which are themselves difficult to pin down in monocausal terminology, exactly because of this multi-causal function, which makes them appear in different circumstances wearing different "masks of appearance."

When considering task versus task comparisons in the case of volume processing, the observed non-significant trend towards a RH involvement is consistent with the majority of the literature on volume processing (Lasota et al. 2003; Opitz et al. 2002; Mustovic et al. 2003, Brancucci et al. 2005). Our results are especially in line with the studies by Belin et al. (1998), who found activation for volume processing within exactly the same region (BA 40) as in our study as being part of a right hemispheric auditory attention network. Since these are only reported trends, we are cautious in interpreting the findings, but would like to suggest that they could be related to the issue of spatial allocation of sounds, since volume is used in distance judgements, which would be a task represented in the dorsal stream (Bushara et al. 1999; Rauschecker and Tian 2000; Warren et al. 2002).

fMRI parametric effects

The parametric effects reported here were only trend activations (see results section). We are therefore reporting the results only as trends and treat the interpretations with caution.

In the case of timbre discrimination (for the linear increase of responses due to accuracy of timbre discrimination, i.e. success rate), we found a trend activation within a right inferior temporal area and the right cerebellum (mainly lobule 6 and 8). The right cerebellar activation could be seen as being connected to and supporting a cortically left-lateralized frontal activation, as is known from the crossed cerebro-cerebellar dominance principle (Jansen et al. 2005). Moreover, the observed performance-dependent activation within the right cerebellum during timbre discrimination (Fig. 4) indicates a cerebellar contribution to this network. Although activation of the cerebellum associated with timbre processing has not been reported so far, the cerebellum has been reported to be involved in a number of basic acoustic processing tasks (Petacchi et al. 2005), mainly related to timing and temporal features (Thaut 2003), as well as language tasks. Related to language, the right cerebellum in particular has been reported to play a role in the representation of speech sound sequences and cognitive tasks that depend upon a phonetic code (Mathiak et al. 2002) and speech perception as well as production, as is the case in auditory verbal

imagery and internal speech which requires the representation of syllabic structure and prearticulatory representation of verbal utterances (Ackermann et al. 2004). It seems plausible that a timbre related, prelinguistic task could be represented within a dynamic network, in which there is a special connection or interplay between Broca's area and the right cerebellum.

Parametric analysis during successful volume discrimination, in contrast, revealed increasing activation within the left lentiform nucleus, culminating in left hippocampal activity. These structures have been shown to be involved in sound distance judgements (Hartley et al. 2004; Kimura et al. 2004). The observed trend of stronger activation in parallel to increasing differences of sound intensities, therefore, is in line with the assumption that volume processing contributes to this process.

Conclusion

Activations within language areas of the brain (left IFG, left AC, right Cerebellum) during processing of non-linguistic acoustic stimuli, indicate that linguistic and non-linguistic processes share resources in the brain and have no strict spatially delineated dedicated areas. This finding is in line with a series of arguments against the existence of macro-anatomical structures dedicated to "speech" based on analysis of functional connectivity patterns during verbal and non-verbal auditory processing (Price et al. 2005).

The observed leftward lateralization within temporal regions during timbre and volume judgements, as well as activation of Broca's area and the right cerebellum during timbre processing, further confirm the involvement and interplay of larger networks, comprising cortical and subcortical structures in "pre-linguistic" acoustic processing. Activation of these networks assumedly depends on the actual task demands and difficulty level (Reiterer et al. 2005), and speaks against a unitary brain area responsible for the processing of timbre or volume.

References

Ackermann, H., Mathiak, K., & Ivry, R. (2004). Temporal organization of "internal speech" as a basis for cerebellar modulation of cognitive functions. *Behavioral and Cognitive Neuroscience Reviews, 3*(1), 14–22.

Belin, P., McAdams, S., Smith, B., Savel, S., Thivard, L., Samson, S., et al (1998). The functional anatomy of sound intensity discrimination. *Journal of Neuroscience, 18*(16), 6388–6394.

Boucher, R., & Bryden, M. P. (1997). Laterality effects in the processing of melody and timbre. *Neuropsychologia, 35*(11), 1467–1473.

Brancucci, A., Babiloni, C., Rossini, P. M., & Romani, G. L. (2005). Right hemisphere specialization for intensity discrimination of musical and speech sounds. *Neuropsychologia, 43*(13), 1916–1923.

Brancucci, A., & San Martini, P. (2003). Hemispheric asymmetries in the perception of rapid (timbral) and slow (nontimbral) amplitude fluctuations of complex tones. *Neuropsychology, 17*(3), 451–457.

Burton, M., Small, S., & Blumstein, S. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience, 12*(4), 679–690.

Bushara, K., Weeks, R., Ishii, K., Catalan, M., Tian, B., Rauschecker, J., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Natural Neuroscience, 2*(8), 759–766.

Deike, S., Gaschler-Markefski, B., Brechmann, A., & Scheich, H. (2004). Auditory stream segregation relying on timbre involves left auditory cortex. *NeuroReport, 15*(9), 1511–1515.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience, 12*(3), 449–460.

Fiez, J., Raichle, M., Miezin, F., Petersen, S., Tallal, P., & Katz, W. (1995). PET studies of auditory and phonological processing: Effects of stimulus characteristics and task demands. *Journal Cognitive Neuroscience, 7*(3), 357–375.

Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Satthamnuwong, N., Tong, Y., et al. (2002). A cross-linguistic FMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience, 14*(7), 1076–1087.

Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society America, 61*, 1270–1277.

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society America, 63*, 1493–1500.

Halpern, R., Zatorre, R., Bouffard, M., & Johnson, J. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*(9), 1281–92.

Hartley, T., Trinkler, I., & Burgess, N. (2004). Geometric determinants of human spatial memory. *Cognition, 94*(1), 39–75.

Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Heim, S., Opitz, B., Muller, K., & Friederici, A. D. (2003). Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Brain Research Cognitive Brain Research, 16*(2), 285–296.

Hertrich, I., & Ackermann, H. (1999). A vowel synthesizer based on formant sinusoids modulated by fundamental frequency. *Journal of the Acoustical Society of America, 106*, 2988–2990.

Hertrich, I., Mathiak, K., Lutzenberger, W., & Ackermann, H. (2003). Processing of dynamic aspects of speech and non-speech stimuli: A whole-head magnetoencephalography study. *Brain Research Cognitive. Brain Research, 17*(1), 130–139.

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain Language, 76*(3), 227–252.

Huang, J., Carr, T., & Cao, Y. (2002). Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping, 15*(1), 39–53.

Iacoboni, M., Molnar-Szakacs, I., Gallese. V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*(3), 79.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*(5449), 2526–2528.

Jäncke, L., Wüstenberg, T., Scheich, H., & Heinze, H. J. (2002). Phonetic perception and the temporal cortex. *NeuroImage, 15*(4), 733–746.

Jansen, A., Floel, A., Van Randenborgh, J., Konrad, C., Rotte, M., Forster, A., et al. (2005). Crossed cerebro-cerebellar language dominance. *Human Brain Mapping, 24*(3):165–172.

Joanisse, M., & Gati, J. (2003). Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. *Neuroimage, 19*(1), 64–79.

Johnsrude, I., Zatorre, R., Milner, B., & Evans, A. (1997). Left hemisphere specialization for the processing of acoustic transients. *NeuroReport, 8*(7), 1761–1765.

Kimura, A., Donishi, T., Okamoto, K., & Tamai, Y. (2004). Efferent connections of "posterodorsal" auditory area in the rat cortex: Implications for auditory spatial processing. *Neuroscience, 128* (2), 399–419.

Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage, 17*(2), 956–966.

Lasota, K., Ulmer, J., Firszt, J., Biswal, B., Daniels, D., & Prost, R. (2003). Intensity-dependent activation of the primary auditory cortex in functional magnetic resonance imaging. *Journal of Computer Assisted Tomography, 27*(2), 213–218.

Levitin, D., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *Neuroimage, 20*(4), 2142–2152.

Mathiak, K., Hertrich, I., Grodd, W., & Ackermann, H. (2002). Cerebellum and speech perception: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience, 14*(6), 902–912.

Menon, V., Levitin, D., Smith, B., Lembke, A., Krasnow, B., Glazer, D., et al. (2002). Neural correlates of timbre change in harmonic sounds. *NeuroImage, 17*(4), 1742–1754.

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*(4), 684–701.

Muller, R. A., Kleinhans, N., & Courchesne, E. (2001). Broca's area and the discrimination of frequency transitions: A functional MRI study. *Brain Language, 76*(1), 70–76.

Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., et al. (2003). Temporal integration of sequential auditory events: Silent period in sound pattern activates human planum temporale. *Neuroimage, 20*(1), 429–434.

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schroger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage, 15*, 167–174.

Petacchi, A., Laird, A., Fox, P., & Bower, J. (2005). Cerebellum and auditory function: An ALE meta-analysis of functional neuro-imaging studies. *Human Brain Mapping, 25*(1), 118–128.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., et al. (1997). The structural components of music perception. A functional anatomical study. *Brain, 120*(2), 229–243.

Price, C., Thierry, G., & Griffiths, T. (2005). Speech-specific auditory processing: Where is it? *Trends in Cognitive Science, 9*(6), 271–276.

Rauschecker, J., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Science of the United States of America, 97*(22), 11800–11806.

Reiterer, S., Erb, M., Droll, C., Anders, S., Ethofer, T., Grodd, W., et al. (2005). Impact of task difficulty on lateralization of pitch and duration discrimination. *Neuroreport, 16*(3), 239–242.

Samson, S. (2003). Neuropsychological studies of musical timbre. *Annals of the New York Acadademy of Science, 999*, 144–151.

Samson, S., & Zatorre, R. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia, 32*(2), 231–240.

Scott, S., & Johnsrude, I. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neuroscience, 26*(2), 100–107.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., et al. (2002). Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study. *Neuroreport, 13*(14), 1813–1816.

Thaut, M. H. (2003). Neural basis of rhythmic timing networks in the human brain. *Annals of the New York Academy of Sciences, 999*, 364–373.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., & Delcroix, N. (2002). Automated anatomical labelling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single subject brain. *Neuroimage, 15*, 273–289.

Vihla, M., & Salmelin, R. (2003). Hemispheric balance in processing attended and non-attended vowels and complex tones. *Brain Research Cognitive Brain Research, 16*(2), 167–173.

Warren, J., Jennings, A., & Griffiths, T. (2005). Analysis of the spectral envelope of sounds by the human brain. *NeuroImage, 24*, 1052–1057.

Warren, J., Zielinski, B., Green, G., Rauschecker, J., & Griffiths, T. (2002). Perception of sound-source motion by the human brain. *Neuron, 34*(1), 139–148.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience, 7*(7), 701–702.

Zatorre, R., Belin, P., & Penhune, V. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science, 6*, 37–46.