Behavioral/Systems/Cognitive

From Air Oscillations to Music and Speech: Functional Magnetic Resonance Imaging Evidence for Fine-Tuned Neural Networks in Audition

Mari Tervaniemi,[1,2]* **André J. Szameitat,**[3]* **Stefanie Kruck,**[4] **Erich Schröger,**[2] **Kai Alter,**[5] **Wouter De Baene,**[6] and **Angela D. Friederici**[4]

[1]Cognitive Brain Research Unit, Department of Psychology, University of Helsinki and Helsinki Brain Research Centre, FIN-00014, Helsinki, Finland, [2]Institut für Psychologie I, Universität Leipzig, 04103 Leipzig, Germany, [3]Department of Psychology, University of Surrey, Guildford GU2 7XH, United Kingdom, [4]Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany, [5]School of Neurology, Neurobiology and Psychiatry, The Medical School, The University of Newcastle, Newcastle NE1 4HH, United Kingdom, and [6]Laboratorium voor Neuro- en Psychofysiologie, Katholieke Universiteit Leuven, Campus Gasthuisberg, 3000 Leuven, Belgium

In the auditory modality, music and speech have high informational and emotional value for human beings. However, the degree of the functional specialization of the cortical and subcortical areas in encoding music and speech sounds is not yet known. We investigated the functional specialization of the human auditory system in processing music and speech by functional magnetic resonance imaging recordings. During recordings, the subjects were presented with saxophone sounds and pseudowords /ba:ba/ with comparable acoustical content. Our data show that areas encoding music and speech sounds differ in the temporal and frontal lobes. Moreover, slight variations in sound pitch and duration activated thalamic structures differentially. However, this was the case with speech sounds only while no such effect was evidenced with music sounds. Thus, our data reveal the existence of a functional specialization of the human brain in accurately representing sound information at both cortical and subcortical areas. They indicate that not only the sound category (speech/music) but also the sound parameter (pitch/duration) can be selectively encoded.

Key words: speech; music; neuroimaging; functional specialization; auditory cortex; thalamus

Introduction

Auditory processing in humans is based on the integrity and complementarity of the left and right temporal lobes. Especially important are the areas in the Heschl's gyrus (HG) as well as in the middle and superior temporal gyri (STG) and sulci (STS) (Tervaniemi and Hugdahl, 2003; Toga and Thompson, 2003; Josse and Tzourio-Mazoyer, 2004; Peretz and Zatorre, 2005). In general, the dominant role of the left hemisphere in speech processing and, to a lesser extent, the role of the right hemisphere in music processing has been highlighted.

Although widely investigated in animal research (Rauschecker et al., 1995; Rauschecker, 1998; Kaas and Hackett, 2000; Tian et al., 2001), functional specialization of the auditory areas has received less attention in human research. However, the rapid development of relevant methodologies has recently enabled investigations into the relationship between advanced auditory

functions and specific temporal-lobe areas. For instance, the bilateral superior temporal sulcus was more strongly activated by speech stimuli than by frequency-modulated tones (Binder et al., 2000) or by nonspeech vocal sounds (Belin et al., 2002; Fecteau et al., 2004). Moreover, sound-change discrimination occurred in a different part of the auditory areas for chord and vowel stimulation (Tervaniemi et al., 1999) and for frequency and duration changes in simple tones (Molholm et al., 2005). However, to accurately map how the human brain represents our every-day sound environment, it is necessary to reveal the neurocognitive architecture in audition by using ecologically valid sound material.

We performed a functional magnetic resonance imaging (fMRI) experiment to separate the auditory areas enabling speech versus music encoding. As speech exemplars we used words rather than sentences because, at the word level, the human brain needs to process information across several hundreds of milliseconds, but without the need to parse high-level syntactic or semantic information. In other words, at the word level, the listener is automatically beyond low-level phonemic processes without a necessary involvement of high-level structural hierarchies know to involve the inferior frontal gyrus (Friederici et al., 2006). We used syllables that are phonotactically legal in German in a two-syllable-pseudoword context. They have the phonological struc-

ture of a real word (e.g., papa/daddy) with stress assigned to the first syllable, which is the predominant stress pattern for two-syllable words in German. In parallel, in music, spectrotemporal sound patterns comparable with (pseudo)words consti-tute already musically meaningful entities as present in every day auditory input. By using sound material of this kind, we wished to probe the neural mechanisms activated by speech and music sound ex-cerpts when used in an experimental par-adigm applied previously with more sim-plified stimulation (Näätänen, 2001; Tervaniemi and Huotilainen, 2003). Dis-tinct neural substrates for encoding speech versus music-sound information in acous-tically comparable inputs would provide strong evidence for within-hemisphere specialization of the auditory cortices in humans.

In addition, we explored the existence of brain areas specific to duration versus frequency encoding. This was motivated by the selective importance of these sound parameters in activating the left versus right hemisphere and by their importance for prosodic processing in speech and emotional processing in music. In other words, the present endeavor examined the existence of sound-type (speech vs music) and parameter-type (duration vs fre-quency) specific tuning of the auditory ar-eas in the healthy human brain by using acoustically complex, ecologically valid sound material.

(a) Sound duration

(b) Sound spectra

(c) Experimental paradigm

Figure 1. *a*, Speech and music sound durations and their relative amplitudes (speech, black; music, gray). *b*, Five formants (speech) and partials (music) between 0 and 5500 Hz. *c*, Schematic illustration of the experimental paradigm.

Materials and Methods

Subjects
There were 17 neurologically healthy adult participants (nine males; mean age, 25.1 years; age range, 22–36) who were paid for their contri-bution. All of them were right-handed according to the Edinburgh In-ventory (Oldfield, 1971) and native German speakers without formal training in music. They had given previous informed consent according to the Max Planck Institute guidelines. The experiment was approved by the ethics review board at the University of Leipzig (Leipzig, Germany).

Stimulation
The scanning session in 3T fMRI consisted of two stimulation condi-tions. In the speech sound condition, the subjects were binaurally pre-sented with a consonant-vowel-consonant-vowel pseudoword /ba:ba/. This was produced by a trained female native speaker of German to preserve the naturalness of human voice in the syllables. In the music sound condition, saxophone sounds with a corresponding acoustical structure were used (Fig. 1*A,B*). These were produced as digital samples from the Montreal timbre library. Both sounds had fundamental fre-quencies between 180 and 190 Hz and were digitized to 44.1 kHz, 16-bit resolution. They were fitted to each other to be similar in envelope con-tour and intensity in the first part of the stimulus.

To increase the informational value of the stimulation, there were three different versions of both sounds, which were perceptually easily discriminable from each other: In addition to the constant "standard" sound, the subjects were presented with two "deviant" sounds with both speech and music sounds. In these, the first part of the stimulus was 75 ms shorter (duration deviant) or 30 Hz higher (frequency deviant) than in standard sounds. Each subject was thus confronted with six different
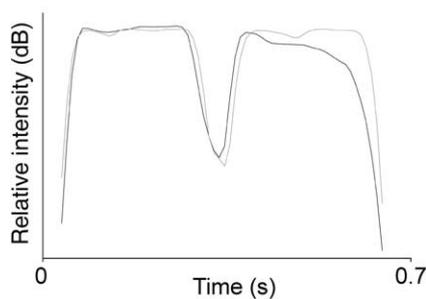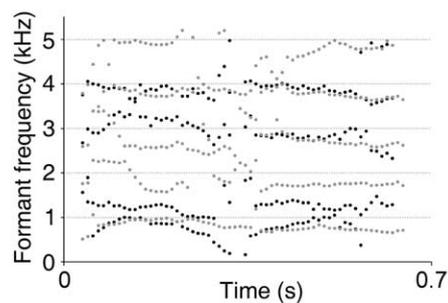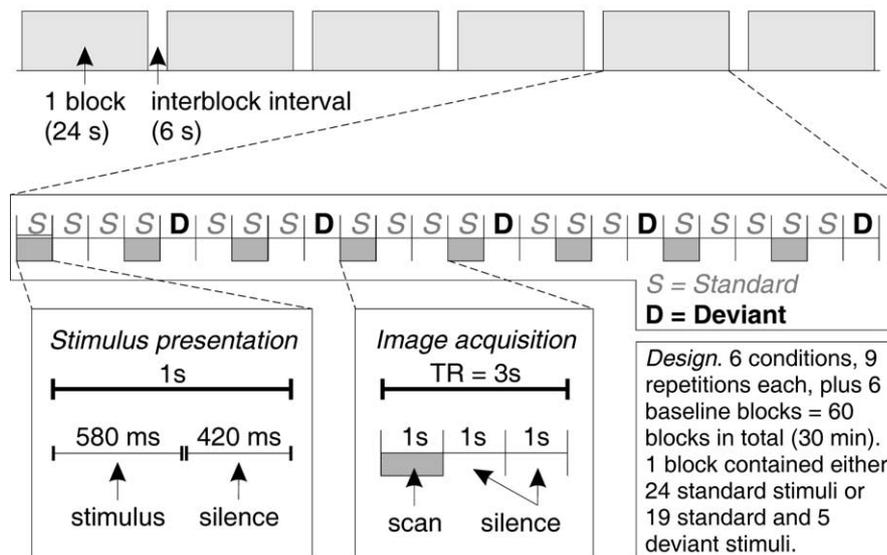
sounds, presented in four different sequences (Fig. 1*C*), with an instruc-tion to indicate whether a given 30 s sequence consisted of music or speech sounds.

fMRI recordings
Procedure. While lying in the fMRI scanner, participants viewed a pro-jection screen via a mirror fixating their eyes on a cross on a screen. The sound stimuli were presented binaurally via headphones of the Com-mander XG system (Resonance Technology, Northridge, CA) with an audible but convenient intensity (70–80 dB, adjusted individually for each subject).

The sounds were presented in blocked design with 1 s stimulus-onset asynchrony. Speech and music sounds were presented in separate blocks (Fig. 1*C*). Each block consisted of 24 sounds of either the original stan-dard sounds alone or, alternatively, of a mixed sequence of standard and duration-deviant sounds or standard and frequency-deviant sounds. In the case of a mixed sequence, there were five deviant sounds presented in a pseudorandom order among the standard sounds. Because of demands set by the clustered scanning paradigm (see below), deviating sounds were never presented in the first, fourth, seventh, etc., position.

The task of the subjects was to indicate whether the sounds were speech or music sounds by pressing a response button with the right index finger for music sounds and the right middle finger for speech sounds. The response had to be performed during a 6 s silent interval, which separated the blocks. The subjects were instructed to pay attention to the sound category of speech and music but they were not informed
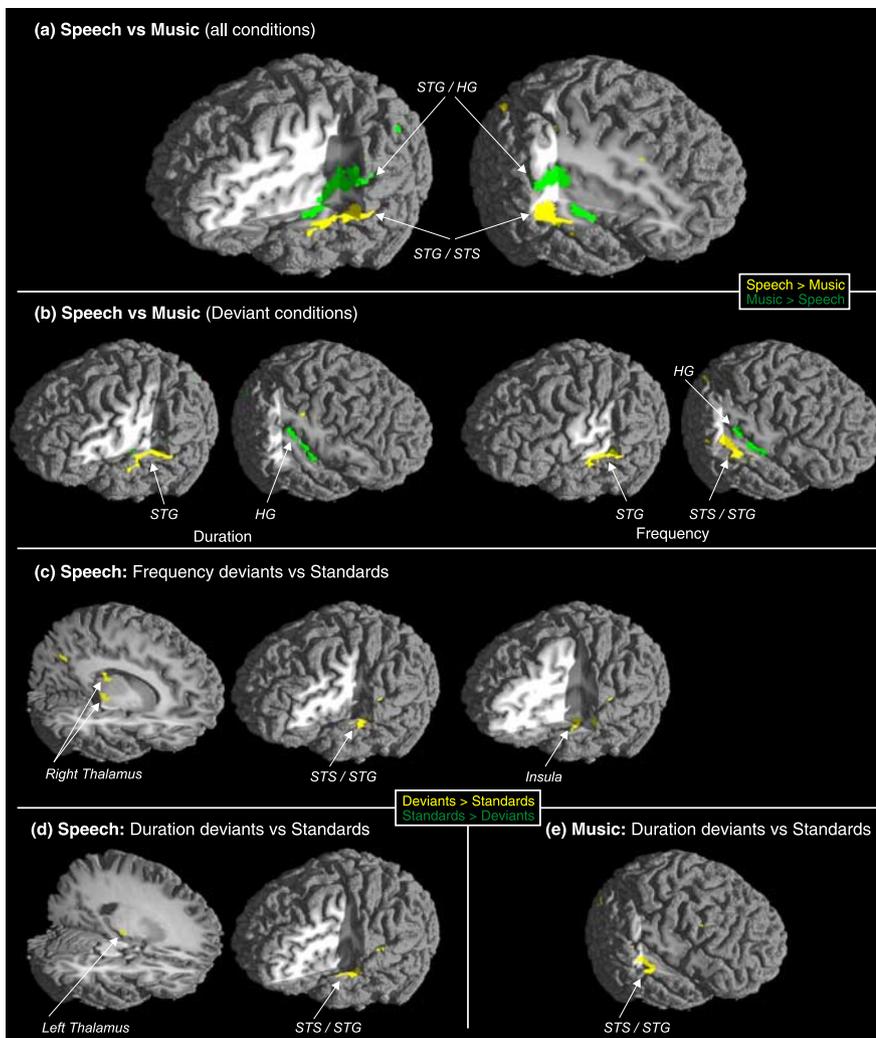
Figure 2. Significant foci of activity in speech versus music contrast (all conditions included) (*a*), speech versus music deviant contrasts (*b*), deviant versus standard contrast for speech (*c, d*), and deviant versus standard for music (*e*). *a, b,* Yellow color illustrates stronger activation for speech than music sounds, and green color illustrates stronger activity for music than speech sounds. *c–e,* Yellow color illustrates stronger activation for deviant than standard sounds, and green color illustrates stronger activity for standard than deviant sounds.

silent period. Before the functional scanning, anatomical scans were obtained (T1 and EPI-T1).

Data analysis. Data were analyzed using the software LIPSIA (Lohmann et al., 2001). After the movement and slice-time corrections, the data of the first 10 volumes of the session were omitted. After the data were temporally filtered (high pass, 0.004 Hz) and spatially smoothed [Gaussian kernel, full-width at half-maximum (FWHM), 5.65 mm], the data of each subject were registered to his/her high-resolution structural T1-image and normalized according to the standard space of Talairach and Tournoux (1988). Thereafter, a three-dimensional data set, rescaled to a voxel size of $3 \times 3 \times 3$ mm^3, was created.

Individual first-level statistical analysis was based on a voxel-wise least squares estimation using the general linear model (GLM) for serially autocorrelated observations (Friston et al., 1995). A boxcar function with a response delay of 6 s was used to generate the design matrix. Design matrix and functional data were linearly smoothed with a 4 s FWHM Gaussian kernel. Next, individual contrast images, containing the β values from the GLM parameter estimation and their variance, were calculated.

Second-level group statistic was based on a Bayesian approach (Neumann and Lohmann, 2003). From the individual contrast images, a posterior probability map (PPM) was calculated, which depicts the probability of a difference between the conditions compared in the contrast of interest. For visualization, PPMs were thresholded with $p > 99.9$ and an extent threshold of 190 mm^3 (Neumann and Lohmann, 2003).

To confirm the results gained by the Bayesian approach, we additionally performed region of interest (ROI) analyses of the peak voxel in all comparisons, which were based on *t* tests according to Fisher statistics. For this, the percent signal change (PSC) in each voxel was individually calculated for each condition and ROI relative to the resting baseline condition. Next, the PSC values were averaged across the blocks, omitting the first and last volume to account for transition effects, resulting in one PSC value for each participant and condition. One-sample repeated measures *t* tests were calculated for each comparison and ROI ($p < 0.05$, two-tailed). Only activation foci showing significant effects in this ROI analysis as well as yielding posterior probabilities of $p > 99.9$ and an extent of 190 mm^3 in the PPMs are reported and discussed.

about the presence of any deviating sounds. During the silent baseline condition, they were to fixate a central cross on the screen.

Each of the six blocks was presented nine times (standards only, duration deviants plus standards, and frequency deviants plus standards, each for speech and music, respectively). The experimental blocks were intermixed with six resting baseline blocks of 30 s each. Additionally, each session was preceded by a 15 s silent baseline to accommodate the subject with the situation. In total, the experiment lasted 30 min per subject.

Data acquisition. The blood oxygen level-dependent (BOLD) responses were recorded by a 3T scanner (Medspec 30/100; Bruker, Ettlingen, Germany), which had a standard birdcage head coil. During the scanning, the participants were lying in a supine position on a scanner bed, cushions being used to prevent head motion.

During functional scanning, 16 axial slices (4 mm thickness, 1 mm interslice gap) parallel to the plane intersecting the anterior and posterior commissures covering the whole brain were acquired by applying a single-shot, gradient recalled echo planar imaging (EPI) sequence [repetition time (TR), 3 s; echo time (TE) 30 ms; 90° flip angle]. The TR was optimized to allow the BOLD signal caused by gradient switching to diminish (although not fully disappear) and to minimize the length of the scanning sessions. Slice acquisition was clustered in the first second of the TR, leaving two seconds without gradient noise (cf. Edmister et al., 1999). For better intelligibility, deviant stimuli occurred only during the

Results

Brain activation: music versus speech sounds

When contrasting all speech sound blocks to all music sound blocks, auditory cortices in the STG were bilaterally active. However, whereas speech sounds activated the inferior part of the lateral STG more than music sounds, music sounds activated the superior/medial surface of the STG/HG more than speech sounds did (Fig. 2*A*, Table 1). Additional frontal activation was observed by speech sounds in the middle frontal gyrus (MFG).

Thus, when the BOLD activity related to neural processing of acoustically complex speech and music sounds is compared, speech sounds activated bilaterally more anterior, inferior, and lateral areas than music sounds. These data on speech sounds correlate closely with recent evidence (Belin et al., 2002; Fecteau

et al., 2004) showing an STS area specialized on human vocalizations when contrasted with animal vocalizations or nonhuman sounds. The present findings obtained with music sounds suggesting relatively more posterior, superior, and medial loci of music information processing within the auditory cortex complement those findings in a novel manner by showing that, despite acoustical similarity, both speech and music sounds have their specialized, spatially distinguishable neural substrates in the human auditory system.

Duration versus frequency deviance in speech versus music sounds

When contrasting speech and music conditions with duration deviants, speech duration deviants activated the left STG more than music duration deviants did (Fig. 2 B, Table 1). In parallel, music duration deviants activated the right HG more than did the speech duration deviants. Thus, the present results on speech duration deviants confirm the lateralization effects in audition by showing left-dominant activation for speech sounds and right-dominant activation for music sounds (Zatorre et al., 2002; Tervaniemi and Hugdahl, 2003).

In a corresponding comparison for frequency deviants, speech frequency deviants activated the right STS/STG, the left STG, and the right MFG more strongly than music frequency deviants. Music frequency deviants activated the right HG more than the speech frequency deviants did. However, frequency deviants produced a more complex pattern of hemispheric lateralization than duration deviants because they activated the right auditory areas in the music domain but bilateral STG areas in the speech domain. We propose that, in addition to the phonemic information activating the left auditory areas, frequency deviants among speech sounds were regarded as prosodic information, thus also activating right auditory areas (for the right hemispheric activation for the processing of sentential prosody, see Meyer et al., 2002). The right MFG activation for the frequency change in speech sounds may imply the involvement of attention- and execution-related neural resources as well. However, because of relatively poor time resolution of the fMRI methodology, the temporal order of activations in the frontal and auditory areas remained open (Opitz et al., 2002; Doeller et al., 2003). Yet, the present finding is in line with the hierarchical arrangement observed between sensory and attentional (frontal) neural systems (Näätänen, 1992; Romanski et al., 1999), with evidence for both bottom-up (Rinne et al., 2000) and top-down projections (Giard et al., 1990; Yago et al., 2001).

Moreover, the areas underlying the deviance processing differed between the music and speech sounds also within the hemispheres: music-dominant activations were more medial in location than their speech-sound equivalents. Thus, auditory areas might in fact constitute high-level specialization both within and between the auditory cortices, as proposed earlier by using more simplified stimulus material (Tervaniemi et al., 1999, 2000; Mol-

Table 1. Activated areas

	Anatomical region	Brodmann area	Talairach stereotactic coordinates of the activation maxima (x, y, z)	Number of activated voxels
Speech versus music contrasts, all blocks				
Speech > music	STG L	BA 22/42	−57, −18, 4	2877
	STG R	BA 22/42	51, −29, 6	4263
	MFG R	BA 8/9	29, 20, 35	212
Music > speech	STG/HG L	BA 22/42, 41	−38, −37, 15	6400
	STG/HG R	BA 22/42, 41	39, −26, 15	5941
Speech versus music contrasts, standard plus deviant blocks				
Speech > music				
Duration	STG L	BA 22/42	−55, −8, 0	1215
Frequency	STG L	BA 22/42	−57, −12, 5	971
	STS/STG R	BA 22/42	51, −27, 4	2853
	MFG R	BA 8/9	23, 53, 21	517
Music > speech				
Duration	HG R	BA 41	35, −25, 15	1318
Frequency	HG R	BA 41	38, −28, 15	1722
Deviant plus standard versus standard contrasts: speech				
Duration > standard	STG/STS L	BA 21/22	−53, −7, 0	380
	Thalamus L		−20, −22, 0	213
Frequency > standard	STS/STG L	BA 22/42	−55, −13, 3	227
	Thalamus R		17, −16, 0	247
	Thalamus R		13, −25, 18	229
	Insula L		−39, −7, −6	358
	Precuneus	BA 7	1, −70, 44	418
Deviant plus standard versus standard contrasts: music				
Duration deviant > standard sounds	STG R	BA 22	50, −20, 1	1070
Frequency deviant > standard sounds	–	–	–	–

BA, Brodmann area; x, left–right; y, posterior–anterior; z, ventral–dorsal.

holm et al., 2005). This line of reasoning is also supported by Patterson et al. (2002), even within the music-sound domain. By using iterated noise stimulation to induce fixed and melodic pitch percepts, they showed the right-dominant auditory areas to constitute a hierarchy in pitch processing, melodic-pitch stimulation activating areas anterior to the HG when contrasted with fixed pitch or noise stimulation.

Discrimination of duration versus frequency deviance in speech versus music sounds

The last set of contrasts for standard and deviant conditions versus standard only conditions revealed the functional organization of memory-based discrimination of duration and frequency changes among standard sounds. In them, standard and deviant sounds together activated the auditory cortex more than standard sound sequences specifically with speech sounds: both frequency and duration deviants activated left STS/STG more than the standard sounds (Fig. 2C,D, Table 1). Moreover, with speech stimuli, these cortical activations were accompanied by parameter-specific lateralized thalamic activations: duration change activated the left and frequency change activated the right thalamic nuclei. In contrast, music sounds with duration deviant were associated with the right-hemispheric STG activation (Fig. 2E, Table 1).

Discussion

Previous fMRI investigations comparing standard and deviant sound processing to reveal memory-based loci of activity were conducted by using sinusoids or harmonic sounds with a few

partials and deviances in pitch and duration. They consistently revealed right STG activation (Opitz et al., 2002; Doeller et al., 2003; Liebenthal et al., 2003; Schall et al., 2003; Sabri et al., 2004; Molholm et al., 2005) and less frequently activation in the left temporal lobe with more variation in the loci of activated sources (Opitz et al., 2002; Liebenthal et al., 2003; Molholm et al., 2005). Frontal sources were also activated when using attention-catching sound changes (Opitz et al., 2002; Doeller et al., 2003; Schall et al., 2003; Molholm et al., 2005). The ecological validity of the previous temporal-lobe findings was upgraded in the present data by using more acoustically complex stimulation with slight deviances. Moreover, the importance of the stimulus familiarity was highlighted in the frontal lobe in which the MFG was activated by the sound material most familiar to any subject without musical expertise, that is, by frequency changes among speech sounds (for language-related activations in corresponding areas, see also Gandour et al., 2003).

Importantly, the present data also indicated subcortical foci of BOLD activity in sound-change discrimination in healthy humans. Until now, in intracranial studies in humans, subcortical sources including the thalamus were activated only when the subjects were instructed to attend to the sounds (Kropotov et al., 2000). In contrast, in guinea pigs, the caudomedial nonprimary portion of the medial geniculate nucleus in the thalamus was activated by pitch-deviant tones (Kraus et al., 1994a) and selectively by the syllabic /ba/-/wa/ but not by /ga/-/da/ contrast (Kraus et al., 1994b). So, the present data highlight the role of the thalamus in memory-based sound processing even in humans, and suggest that feature-specific processes include subcortical specialization and can be followed up to the auditory cortices. Notably, this result was obtained with the stimuli with which the subjects were highly familiar, namely, speech sounds.

During the fMRI scanning, the subjects were categorizing the sound sequences into speech or music sounds without being informed about the presence of any sound changes. Our data, thus, index functional specialization of the auditory areas to differentially encode speech and music material during attentional listening (Fig. 2a, b). However, the sound-change related activations reflect functioning of neural mechanisms even without focused attention because the sound changes were not relevant to the task performance (Fig. 2c–e).

Together, the present results provide evidence that in the future, the functional architecture of human audition can be mapped cortically and subcortically with the precision currently available in nonhuman primates (Rauschecker et al., 1995; Rauschecker, 1998; Kaas and Hackett, 2000; Tian et al., 2001). By these means, knowledge about the brain's basis for human cognition and emotion in the auditory domain will expand, potentially giving rise to beneficial and well targeted applications in education and therapy.

References

Belin P, Zatorre RJ, Ahad P (2002) Human temporal-lobe response to vocal sounds. Cognit Brain Res 13:17–26.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activations by speech and nonspeech sounds. Cereb Cortex 10:512–528.

Doeller CF, Opitz B, Mecklinger A, Krick C, Reith W, Schröger E (2003) Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. NeuroImage 20:1270–1282.

Edmister WB, Talavage TM, Ledden PJ, Weisskoff RM (1999) Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 7:89–97.

Fecteau S, Armony JL, Joanette Y, Belin P (2004) Is voice processing species-

specific in human auditory cortex? An fMRI study. NeuroImage 23:840–848.

Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A (2006) The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc Natl Acad Sci USA 103:2458–2463.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2:189–210.

Gandour J, Xu Y, Wong D, Dzemidzic M, Lowe M, Li X, Tong Y (2003) Neural correlates of segmental and tonal information in speech perception. Hum Brain Mapp 20:185–200.

Giard MH, Perrin F, Pernier J, Bouchet P (1990) Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. Psychophysiology 27:627–640.

Josse G, Tzourio-Mazoyer N (2004) Hemispheric specialization for language. Brain Res Rev 44:1–12.

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA 97:11793–11799.

Kraus N, McGee T, Littman T, Nicol T, King C (1994a) Nonprimary auditory thalamic representation of acoustic change. J Neurophysiol 72:1270–1277.

Kraus N, McGee T, Carrell T, King C, Littman T, Nicol T (1994b) Discrimination of speech-like contrasts in the auditory thalamus and cortex. J Acoust Soc Am 96:2758–2768.

Kropotov JD, Alho K, Näätänen R, Ponomarev VA, Kropotova OV, Anichkov AD, Nechaev VB (2000) Human auditory-cortex mechanisms of preattentive sound discrimination. Neurosci Lett 280:87–90.

Liebenthal E, Ellingson ML, Spanaki MV, Prieto TE, Ropella KM, Binder JR (2003) Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. NeuroImage 19:1395–1404.

Lohmann G, Müller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001) LIPSIA–a new software system for the evaluation of functional magnetic resonance images of the human brain. Comput Med Imaging Graph 25:449–457.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY (2002) fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp 17:73–88.

Molholm M, Martinez A, Ritter W, Javitt DC, Foxe JJ (2005) The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. Cereb Cortex 15:545–551.

Näätänen R (1992) Attention and brain function. Hillsdale, NJ: Erlbaum.

Näätänen R (2001) The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent MMNm. Psychophysiology 38:1–21.

Neumann J, Lohmann G (2003) Bayesian second-level analysis of functional magnetic resonance images. NeuroImage 20:1346–1355.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E (2002) Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. NeuroImage 15:167–174.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. Neuron 36:767–776.

Peretz I, Zatorre RJ (2005) Brain organization for music processing. Annu Rev Psychol 56:89–114.

Rauschecker JP (1998) Parallel processing in the auditory cortex of primates. Audiol Neurootol 3:86–103.

Rauschecker JP, Tian B, Hauserm M (1995) Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268:111–114.

Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Näätänen R (2000) Separate time behaviors of the temporal and frontal mismatch negativity sources. NeuroImage 12:14–19.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136.

Sabri M, Kareken DA, Dzemidzic M, Lowe MJ, Melara RD (2004) Neural correlates of auditory sensory memory and automatic change detection. NeuroImage 21:69–74.

Schall U, Johnston P, Todd J, Ward PB, Michie PT (2003) Functional neu-

roanatomy of auditory mismatch processing: an event-related fMRI study of duration-deviant oddballs. NeuroImage 20:729–736.

Talairach J, Tournoux P (1988) Co-planar stereotactic atlas of the human brain. New York: Thieme.

Tervaniemi M, Hugdahl K (2003) Lateralization of auditory-cortex functions. Brain Res Rev 43:231–246.

Tervaniemi M, Huotilainen M (2003) The promises of change-related brain potentials in cognitive neuroscience of music. Ann NY Acad Sci 999:29–39.

Tervaniemi M, Kujala A, Alho K, Virtanen J, Ilmoniemi RJ, Näätänen R (1999) Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. NeuroImage 9:330–336.

Tervaniemi M, Medvedev SV, Alho K, Pakhomov SV, Roudas MS, van Zuijen TL, Näätänen R (2000) Lateralized automatic auditory processing of phonetic versus musical information: a PET study. Hum Brain Mapp 10:74–79.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. Science 292:290–293.

Toga AW, Thompson PM (2003) Mapping brain asymmetry. Nat Rev Neurosci 4:37–48.

Yago E, Escera C, Alho K, Giard MH (2001) Cerebral mechanisms underlying orienting of attention towards auditory frequency changes. NeuroReport 12: 2583–2587.

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6:37–46.