We aimed at testing the cortical representation of complex natural sounds within auditory cortex using human functional magnetic resonance imaging (fMRI). To this end, we employed 2 different paradigms in the same subjects: a block-design experiment was to provide a localization of areas involved in the processing of animal vocalizations, whereas an event-related fMRI adaptation experiment was to characterize the representation of animal vocalizations in the auditory cortex. During the first experiment, we presented subjects with recognizable and degraded animal vocalizations. We observed significantly stronger fMRI responses for animal vocalizations compared with the degraded stimuli along the bilateral superior temporal gyrus (STG). In the second experiment, we employed an event-related fMRI adaptation paradigm in which pairs of auditory stimuli were presented in 4 different conditions: 1) 2 identical animal vocalizations, 2) 2 different animal vocalizations, 3) an animal vocalization and its degraded control, and 4) an animal vocalization and a degraded control of a different sound. We observed significant fMRI adaptation effects within the left STG. Our data thus suggest that complex sounds such as animal vocalizations are represented in putatively nonprimary auditory cortex in the left STG. Their representation is probably based on their spectrotemporal dynamics rather than simple spectral features.

Keywords: complex sounds, fMRI adaptation, natural sounds, sound recognition

Introduction

The ability to quickly identify complex sounds enables us to interact adequately with our environment when visual information is scarce. Neuropsychological studies have described brain lesions that lead to impaired processing of complex auditory sounds (Engelien et al. 1995; Clarke et al. 2000, 2002). This deficit—auditory agnosia—has been associated with left, right, or bilateral lesions of temporal or temporoparietal cortex, suggesting specialized cortical networks for complex sound processing. However, the coarse localization of lesions has prevented the identification of the exact anatomical correlates of higher order auditory processing. Neurophysiological studies in nonhuman primates provided evidence for hierarchical processing in macaque primary and nonprimary auditory cortex (Rauschecker et al. 1995; Kaas et al. 1999; Rauschecker and Tian 2000; Tian et al. 2001). More specifically, histological studies described a subparcellation of macaque auditory cortex into core and adjacent belt regions (Morel et al. 1993; Jones et al. 1995; Hackett et al. 1998). Functionally, neurons in the core areas have been shown to respond well to pure tones (Rauschecker et al. 1997), whereas belt areas show a preference for more complex stimuli such as band-passed noise (Rauschecker et al. 1995). Moreover, neural populations in lateral belt areas appear to be selective for both the rate and the direction of frequency modulation (Rauschecker 1997, 1998a, 1998b) and for monkey vocalizations (Rauschecker and Tian 2000). For human subjects, a similar hierarchy of auditory processing has been reported, but the actual homologies are still under debate. In particular, anatomical and histochemical studies have described a subdivision into core, belt, and parabelt auditory cortex in the superior temporal lobe; homologous to the macaque anatomy (Sweet et al. 2005). Recent functional magnetic resonance imaging (fMRI) studies have revealed the existence of at least 2 tonotopically organized areas on the superior temporal plane, apparently corresponding to macaque A1 and R (Formisano et al. 2003). Furthermore, a central core area showing tonotopic organization is surrounded by belt areas that exhibited stronger fMRI responses to band-pass noise compared with pure tones (Wessinger et al. 2001; Seifritz et al. 2006). For nonprimary auditory cortex lateral and posterior to Heschl’s gyrus (HG), recent fMRI studies have observed stronger responses to amplitude and frequency modulated tones compared with unmodulated tones (Hall et al. 2002; Hart et al. 2003). Interestingly, the areas that exhibit the 2 types of modulation are largely overlapping, suggesting common neural correlates. Both amplitude and frequency modulation are properties of complex environmental sounds that might serve as basic features to process these sounds. A series of fMRI studies presented subjects with complex natural sounds (animal voices, tools, dropped objects, and liquids) and compared the fMRI responses for recognizable sounds with unrecognizable reversed sounds (Lewis et al. 2004). The findings of these studies suggested that a predominantly left-hemispheric semantic network is involved in the processing of recognizable complex sounds, in addition to bilateral foci in the posterior portions of the middle temporal gyrus. When different categories of natural sounds were compared, stronger fMRI activity was found in a predominantly left-lateralized cortical network including the left and right posterior portions of the middle temporal gyrus for tools versus animal sounds (Lewis et al. 2005). In contrast, the bilateral middle portion of the superior temporal gyrus (mSTG) exhibited a preference for animal compared with tool sounds.

Another important function of the human auditory system is the processing of speech sounds. For example, stronger fMRI responses to speech compared with frequency-modulated tones have been demonstrated in the superior temporal sulcus (STS) (Binder et al. 2000). However, in the same study, similar fMRI responses were obtained for reversed speech and pseudowords, suggesting that these areas are responsive to the acoustic properties of speech rather than language per se. Interestingly, results from an fMRI study that compared cortical responses
with consonant-vowel syllables to pure tones and noise have suggested an implication of bilateral STS and the planum temporale (PT) in language-specific processes (Jäncke et al. 2002). In line with this evidence, by employing sine-wave speech stimuli, a recent fMRI study has shown involvement of the left posterior STS in speech recognition (Möttönen et al. 2006).

The comparison of different categories is able to reveal brain regions with differential preferences for one category over the other. However, this approach may not be appropriate to identify areas that are involved in the processing of more than one category. In particular, an area that has spatially overlapping or close neural populations encoding different stimulus categories would not show differential fMRI responses due to the limited spatial resolution of fMRI and would thus be regarded as not involved in the encoding of the respective categories. To overcome this problem, the present study employed an fMRI adaptation paradigm, which takes advantage of the observation that fMRI responses decrease with repeated presentation of the same stimuli. This technique has been widely applied in the investigation of the human visual system (for a review, Grill-Spector et al. 2006). For example, a recent fMRI adaptation study (Altmann et al. 2004) has tested the effect of visual context on shape processing in higher visual areas. For the auditory domain, the evidence for fMRI adaptation effects is rare. For instance, adaptation effects for speaker’s voice have been demonstrated in the anterior part of the right STS (Belin and Zatorre 2003).

Furthermore, a recent study has described correlations between behavioral effects of repetition priming and repetition-associated reduction of fMRI responses to environmental sounds in the right STS, the bilateral STSs and right inferior frontal gyrus (Bergerbest et al. 2004). Previous studies that tested for repetition priming effects in the auditory domain were not able to reveal significant reductions in the fMRI signal (Badgaiyan et al. 1999; Buckner et al. 2000; Wheeler et al. 2000; Badgaiyan et al. 2001). Bergerbest et al. (2004) argued that this lack of repetition-related reductions could possibly be explained by the fact that these studies have used auditory word-priming tasks that rely more on phonological rather than acoustic representations. For this reason, it might be more adequate to employ environmental sounds rather than spoken words to test for repetition effects in the auditory cortex.

As the processing of different categories of auditory objects differs in its spatial (Lewis et al. 2005) and temporal (Murray, Camen et al. 2006) properties, we decided to test fMRI adaptation effects for a single category of complex sounds only, that is, animal vocalizations. This class of environmental sounds is usually experienced and learned early in child development and refers to tangible and perceptually rich objects in our environment. Thus, the present study used human fMRI to investigate the neural mechanisms that underlie the representation of animal vocalizations as a specific category of complex natural sounds. To this end, we employed 2 different techniques: first, we conducted a block-design experiment for a coarse localization of cortical areas that are involved in the processing of animal vocalizations. Second, we employed an fMRI adaptation paradigm to test for the form of representation of animal vocalizations within the auditory cortex. In the block-design experiment, we presented subjects with animal vocalizations and control sounds with similar spectral features but distorted temporal structure. We manipulated the degree to which the animal vocalizations were degraded and tested for a correlation between the degree of scrambling with blood oxygen level-dependent responses in the human auditory cortex. In the fMRI adaptation paradigm, we tested for selective representation of animal vocalizations. To this end, we presented pairs of sounds that were 1) 2 identical animal vocalizations, 2) 2 different animal vocalizations, 3) an animal vocalization and its degraded control, and 4) an animal vocalization and a degraded control of a different animal vocalization. Brain areas that selectively represent animal vocalizations should exhibit stronger fMRI responses for different compared with same animal vocalizations. In contrast, areas that mainly represent simple spectral features should exhibit cross-adaptation when an animal vocalization and its degraded control is presented and a rebound from adaptation in response to the combination of an animal vocalization and a degraded control of a different animal vocalization.

Materials and Methods

Subjects
Thirteen healthy, right-handed volunteers (age range 19–38, 8 male) participated in the 2 experiments. Handedness was evaluated by self-report. All subjects had normal hearing abilities and gave their informed consent to participate in the study. The experiments were performed in accordance with the ethical standards laid down in the 1991 declaration of Helsinki.

Stimuli
Ten different animal vocalizations (cat, chicken, cow, dog, donkey, duck, frog, horse, owl, and sheep) were taken from a database specifically designed for auditory psychophysics (Marcell et al. 2000). Sounds were digitized with a sampling rate of 44 100 Hz. Sound duration was on average 990 ms (standard deviation 180 ms, range: 650–1250 ms), sound intensity was individually adjusted to a comfortable level resulting in an average sound pressure level of 86 dB. The sounds were equalized as regards their root mean square energy and presented with MR-compatible headphones (Resonance Technology, Northridge, CA). The frequency response of the headphones ensured reliable presentation of the employed stimuli. More specifically, the frequency response was maximal at about 1.5 kHz (99 dBC), remaining at about 82-88 dB for lower frequencies (0.25–1 kHz), and at about 83-88 dB for higher frequencies (2–8 kHz).

We verified the recognizability of the animal vocalizations in a separate session, instructing 6 naive subjects to spontaneously name each animal. We employed the same auditory stimulus system as used in our fMRI experiments. The naming performance showed that the animals were reliably recognized (cat: 6/6; chicken: 6/6; cow: 6/6; dog: 6/6; donkey: 6/6; duck: 3/6; frog: 6/6; horse: 6/6; owl: 6/6; and sheep: 6/6).

As shown in Figure 1, control stimuli consisted of degraded versions of the animal vocalizations that were generated in the following way: we performed a short-term Fourier transform on each stimulus (windows of 64 time points) and increasingly degraded the animal vocalizations by permuting the values of the amplitude and phase components across time (0%, 25%, 50%, 75%, and 100% degradation). Then, we performed the inverse discrete Fourier transform on that signal to recover the amplitude waveform (Fig. 1). The resulting stimuli preserved the average output frequency and similar overall spectral composition. The temporal structure, however, namely the amplitude envelope, and periodicity was disrupted. Previous fMRI studies employed a similar approach, permuting the amplitude and phase values across frequencies, preserving the overall temporal structure (Belin et al. 2002), or generating spectrally matched noise (Thierry et al. 2003) with disrupted temporal structure.

Procedure
All subjects were administered 2 runs of the block-design experiment, 4 runs of the event-related fMRI adaptation experiment and a high-resolution T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence as an anatomical reference. The block-design experiment consisted of 15 blocks that lasted 30 s each with a 15-s stimulation period and 15-s scanner noise only period without...
additional auditory stimulation (baseline). Within each stimulation block, 10 sounds were presented with an interstimulus interval of 1.5 s. Five different conditions that is 5 different degrees of degradation (0%, 25%, 50%, 75%, and 100% degraded) were presented blockwise. During the full run that lasted 7 min and 40 s, subjects were instructed to keep their eyes open and fixate a white cross on a black background, back projected by a LCD-projector through a mirror fixed to the head coil. Within each block, one single stimulus was repeated in sequence. Subjects were instructed to respond to this repetition with a button press to ensure attentive processing of the stimuli.

For the event-related fMRI experiment, an experimental run comprised 80 pairs of auditory stimuli. Each experimental run had a duration of 8 min 30 s and contained 4 different conditions: 1) 2 identical animal vocalizations, 2) 2 different animal vocalizations, 3) an animal vocalization and its degraded control (100% degraded), and 4) an animal vocalization and a degraded control (100% degraded) of a different animal vocalization. Within a single run, 20 trials per condition were presented and intermixed with 20 trials without auditory stimulation (baseline). Similar to previous studies (Altmann et al. 2004), the order of presentation was counterbalanced so that trials from each condition were preceded (1 trial back) equally often by trials from each of the other conditions. The intertrial interval was 5 s and the delay between first and second stimulus was 510 ms on average. To ensure efficient parameter estimation of the hemodynamic response function, the onset of the first stimulus was jittered by ±500 ms (Dale et al. 1999). During the whole experiment, subjects were instructed to keep their eyes open and look at a fixation cross. Subjects were required to press a button with one hand when a stimulus pair contained a degraded control stimulus and a different button with the other hand when a stimulus pair contained 2 animal vocalizations. Subjects showed above 98% correct responses in all conditions at this task.

**Imaging and Data Analysis**

Imaging was conducted on a 3T Siemens ALLEGRA scanner at the Brain Imaging Center Frankfurt, Germany. An echo planar imaging sequence with gradient echo sampling (repetition time [TR] = 2 s, echo time [TE] = 35 ms) was used to acquire the functional imaging data.Thirty-four axial slices (3 mm thickness with 3.00 × 3.00 mm in-plane resolution, interslice gap: 0.3 mm), covering the whole brain, were collected with a head coil. The field of view was 19.2 × 19.2 cm with an in-plane resolution of 64 × 64 pixels. As anatomical reference, we acquired 3-dimensional (3D) volume scans by using a MPRAGE sequence (160 slices; TR: 2300 ms).

fMRI data were preprocessed and analyzed using the Brainvoyager QX (Brain Innovation, Maastricht, Netherlands) software package. Preprocessing of functional data included slice scan time correction, head movement correction, and linear detrending. The 2D functional images were coregistered with the 3D anatomical data, and both 2D and 3D data were spatially normalized into the Talairach and Tournoux stereotactic coordinate system (Talairach and Tournoux 1988). Functional 2D data were spatially filtered employing a Gaussian filter with 4-mm full width half maximum for the group analysis only. The experimental effects in the different conditions were compared employing a random-effects general linear model. To display statistically significant differences, we employed a criterion of $P < 0.001$, uncorrected for multiple comparisons. This criterion was chosen to be equal to the constant criterion used in the single-subject regions of interest (ROI) selection for experiment 2. To achieve anatomical correspondence and to reduce the intersubject variability of cortical gyification, we applied cortex-based intersubject alignment (Fischl et al. 1999).

For the event-related fMRI adaptation experiment, we selected 3 ROIs for each individual subject. An area within the STG was selected functionally on the basis of experiment 1. More specifically, for each individual subject, we selected all voxels within the bilateral STG that showed a significant decrease of the fMRI responses ($P < 0.001$ uncorrected) with the degradation of the animal vocalizations but significantly stronger responses for all sounds compared with the baseline stimulation in experiment 1. HG and the PT were determined in a first step anatomically, employing the anatomical landmarks of the cortical surface of each individual subject (Brechmann et al. 2002). Within the anatomically defined areas, we selected functionally those areas that showed significantly stronger fMRI responses for all auditory

**Figure 1.** Sample stimulus (duck vocalization). Amplitude waveform (upper panel) of the original stimulus (left), 50% degraded (middle), and 100% degraded (right). Spectrograms (lower panel) of the original stimulus (left), 50% degraded (middle), and 100% degraded (right).
stimuli of experiment 1 compared with the baseline. This procedure aimed at avoiding averaging across neural populations that are tuned to frequencies beyond the range of the employed stimuli. Finally, our ROI analysis in HG and PT was restricted to those voxels that fell within the 90% boundaries of the probability map for HG (Rademacher et al. 2001) and the 95% boundaries for PT (Westbury et al. 1999), respectively. All ROIs were selected not to overlap with each other in order to avoid dependencies between them. For all ROIs, we conducted a general linear model to estimate the beta weights that model the hemodynamic responses as a combination of 2 gamma functions within these areas (Friston et al. 1995, 1998). The beta weights were related to the overall signal level to express them as percent signal change values. In Figure 3, the standard error bars of the fMRI responses were corrected for intersubject effects similar to previous studies (Altmann et al. 2004). In particular, the fMRI responses were calculated individually for each subject by subtracting the mean percent signal change for all the conditions within this subject from the mean percent signal change for each condition and adding the mean percent signal change for all the conditions across subjects. We then compared the percent signal change values of all subjects across conditions by a one-way analysis of variance (ANOVA) with repeated measurements. The degrees of freedom were corrected with the Greenhouse-Geisser method when appropriate. In case of a significant ANOVA result, we conducted further linear contrast analysis to compare fMRI responses to same and different animal vocalizations and to compare pairs with control stimuli generated from same or different animal vocalizations. This analysis aimed at characterizing the adaptation effects for animal vocalizations and to test for cross-adaptation from an animal vocalization to its control.

**Results**

In experiment 1, we tested the neural correlates of complex sound processing by investigating the relationship between the parametrically manipulated degradation of animal vocalizations and the acquired fMRI responses. We employed a random-effects general linear model to characterize common fMRI responses to animals versus degraded control stimuli across the subjects (n = 15). As shown in Figure 2a, linear contrast analysis revealed bilateral areas in the STG that exhibited a significant (P < 0.001 uncorrected) decrease of fMRI responses with degradation of animal vocalizations. As depicted on Figure 2b, these effects were robustly observed within individual subjects.

To assess the selective representation of animal vocalizations versus control stimuli with similar spectral features, we conducted an event-related fMRI adaptation experiment. We employed ROI analysis to characterize the fMRI response profile within the auditory cortex for this experiment. The ROIs were based on experiment 1; ROI sizes and locations are shown in Table 1. Left and right HG did not differ significantly in size (paired 2-tailed t-test: t(12) = 1.92; P = 0.08), similar to left and right STG (t(12) = 1.38; P = 0.19). However, the left PT ROI was significantly larger than right PT (t(12) = 3.78; P < 0.001).

As shown in Figure 3, within the areas in the STG that exhibited significantly stronger fMRI responses for animal vocalizations versus degraded sounds, we observed significant differences across experimental conditions for the adaptation experiment. In particular, a repeated measurements ANOVA on the percent signal change levels as shown in Figure 3B with the factor condition (same animal vocalizations, different animal vocalizations, animal—same control, animal—different control) revealed significantly different fMRI responses across conditions in the left STG (F1,12 = 12.89; P < 0.001) but not for the right STG (F1,12 = 2.59; P = 0.07). A linear contrast analysis in the left STG showed significantly stronger fMRI responses for different versus same animal vocalizations (F1,12 = 9.70; P < 0.01), indicating a significant rebound from adaptation. The interaction between the factors hemisphere (left/right STG) and repetition (same/different animal vocalization) was significant (F1,12 = 12.39; P < 0.05), suggesting differential processing across the 2 hemispheres.

We furthermore compared conditions with animal vocalizations only and pairs of animal vocalizations and control stimuli. A repeated measurements ANOVA revealed significantly lower fMRI responses for the pairs that contained control stimuli within bilateral STG (left STG: F1,12 = 20.05, P < 0.001; right STG: F1,12 = 6.90, P < 0.05). Because the ROIs were selected on the basis of the previous block-design experiment as areas with stronger responses for animal vocalizations versus control stimuli, it was expected that they would also show this effect in the fMRI adaptation experiment. However, if these ROIs were selective for simple spectral features rather than the temporal structure of stimuli, we should also expect fMRI adaptation effects for the conditions when the animal vocalizations are paired with their degraded versions compared with different control stimuli. We did not observe significant differences between same or different control stimuli in the STG (left hemisphere: F1,12 < 1, P = 0.56; right hemisphere: F1,12 < 1, P = 0.66). Apart from this functionally defined ROI, we also
analyzed the fMRI responses of the medial portion of HG and the PT. We observed significant differences between experimental conditions in the right PT only ($F_{3.36} = 5.38; P < 0.01$). More specifically, pairs with control stimuli showed a significantly stronger response compared with pairs of animal vocalizations only as determined by a 2-way repeated measurements ANOVA ($F_{1,12} = 9.87; P < 0.01$). Within the right PT, no adaptation effects were revealed for different versus same animal vocalizations ($F_{1,12} < 1; P = 0.45$) or pairs of animal vocalizations and same versus different control ($F_{1,12} < 1; P = 0.98$). As the fMRI responses for control stimuli were not larger compared with the animal vocalizations in experiment 1 (paired t-test: $t_{12} = 0.94; P = 0.37$), the differences in experiment 2 are possibly based on the stimulus change, rather than the acoustic properties of single stimuli. We did not observe any significant experimental effects in HG or left PT (left HG: $F_{3.36} = 1.24; P = 0.30$; right HG: $F_{3.36} = 1.53; P = 0.22$; left PT: $F_{3.36} = 1.88; P = 0.15$). There was no significant interaction between the factors hemisphere (left/right PT) and stimulus type (animal vocalizations only/animal vocalization and control) as tested with a 2-way repeated measurements ANOVA ($F_{1,12} = 2.92; P = 0.11$).

**Discussion**

Employing 2 different paradigms, our study provides evidence for processing of animal vocalizations in the bilateral STG. First, regions along bilateral STG showed a linear decrease of fMRI response strength with the degree of degradation of animal vocalizations. Second, within the left STG, fMRI adaptation effects were observed for same versus different animal vocalizations, suggesting a selective representation of these stimuli.

The bilateral pattern of fMRI responses to animal vocalizations is in line with studies that observed preferential fMRI responses to animal sounds versus tool sounds (Lewis et al. 2005) within the mSTG. Accordingly, areas within bilateral mSTG and the STS have been implicated in the processing of human voices and vocalizations (Belin et al. 2000; Fecteau et al. 2004, 2005; Lewis et al. 2004). Furthermore, among other areas, the STG has been shown to respond more strongly to environmental sounds compared with white noise stimuli (Giraud and Price 2001; Maeder et al. 2001). In a recent study, fMRI responses in the mSTG exhibited an increase for both recognized and unrecognized environmental sounds (Lewis et al. 2005). Thus, it has been proposed that the bilateral mSTG might represent an "intermediate" stage of auditory processing. That is, these areas are possibly involved in the encoding of acoustic structural attributes such as harmonic or phase-coupling content rather than higher level semantic information. In line with this assumption, areas in the superior temporal lobe have been associated with processing of temporal and spectral aspects of auditory stimuli. For example, recent fMRI studies have revealed overlapping areas within the lateral HG and the PT that are sensitive to both amplitude and frequency-modulated tones (Giraud et al. 2000; Hart et al. 2003). Intracranial recordings in human patients have demonstrated encoding of amplitude modulation in the auditory cortex (Ligeois-Chauvel et al. 2004). Moreover, previous electrophysiological studies have suggested directional selectivity of neurons for frequency-modulated sweeps in posterior nonprimary auditory fields of the cat (Tian and Rauschecker 1998). Possibly, the selective representation of the spectrotemporal dynamics of auditory stimuli constitutes the basis for the decoding of complex natural sounds such as animal vocalizations. Thus, similar adaptation effects within the STG might be expected for other stimulus classes with complex spectrotemporal dynamics such as human verbal and nonverbal vocalizations or other environmental sounds.

In this study, the bilateral STG exhibited stronger fMRI responses for animal vocalizations compared with temporally unstructured control stimuli. Adaptation effects, however, occurred within the left STG only. A possible explanation for the lateralization of adaptation effects might be the differential preference for timing parameters within left and right auditory cortex that has been reported in several neuroimaging studies. More specifically, the right STS has been shown to respond preferentially to slowly modulated signals, whereas the left STS showed preferences for faster spectrotemporal changes (Belin

**Figure 3.** (a) Event-related time courses of fMRI responses in the left and right superior temporal gyrus (STG) for experiment 2 averaged across 13 subjects. (b) fMRI response amplitudes in HG, the PT, and the STG for experiment 2 averaged across 13 subjects. The error bars represent mean standard errors of the fMRI responses.
et al. 1998; Boemio et al. 2005). Similarly, within the human anterolateral auditory belt, temporal auditory processing has been associated with the left hemisphere and spectral processing with the right hemisphere (Zatorre and Belin 2001; Schönwiesner et al. 2005). Thus, the observation of mainly left-hemispheric adaptation effects might be accounted for by the specific spectrotemporal dynamics of the employed animal vocalizations. That is, the spectrotemporal changes that distinguish between these stimuli possibly lie within a time range that is preferentially processed by the left STG. A previous study on fMRI repetition priming showed bilateral repetition-related reductions in the STS but right-lateralized reductions within the STG (Bergerbest et al. 2004). In contrast to the present experiments, this study employed general environmental sounds (animals and sounds produced by man-made objects) that might explain the obvious different hemispheric lateralization of repetition effects. Interestingly, a previous magnetoencephalographic (MEG) study has shown a bilateral distribution of evoked magnetic fields in response to mismatch of complex sounds (syllables, animal vocalizations, and noise), but a left-hemispheric increase in gamma-band activity above 60 Hz (Kaiser et al. 2002; Kaiser and Lutzenberger 2005). Induced gamma-band activity has been associated with repetition priming effects in electroencephalographic studies (Gruber and Müller 2002, 2005) and has been suggested to play a complementary role compared with evoked potentials. Thus, the previously reported left-lateralized increase in gamma-band activity in response to pattern changes of complex sounds may be related to the left-lateralized adaptation effects observed in this study.

A different theory of auditory hemispheric lateralization holds that the left superior temporal lobe is preferentially involved in speech perception, whereas the right superior temporal lobe is more involved in environmental sound processing (Thierry et al. 2003). However, a recent review argued that processing of speech and environmental sounds often implicates overlapping areas, concluding that there are no speech-specific cortical regions, but that common correlates might mediate both processes (Price et al. 2005). Thus, the areas in the bilateral STG may be part of a multipurpose processing stage for complex sounds of which the left STG selectively encodes the stimulus features that optimally distinguish between the different animal vocalizations used in the present study.

Interestingly, we observed significantly stronger fMRI responses in right PT for changes from animal vocalizations to degraded control stimuli compared with pairs of animal vocalizations only. In experiment 1, we did not observe significant differences between stimuli, suggesting that the effects within right PT are due to the stimulus change rather than the properties of single stimuli. Possibly, in line with the hypothesis that lateralization of auditory processing is based on preferential tuning to temporal properties, right PT may be involved in the representation of spectrotemporal stimulus features with a long time constant. More specifically, a change from an animal vocalization to the control stimulus is a transition from a temporally structured stimulus to a sustained presentation of a broadband stimulus. Such a transition might possibly be encoded in areas that preferentially respond to slower modulations of auditory stimuli. Nevertheless, future research is required to test systematically for effects of timing parameters of complex auditory stimuli on auditory fMRI adaptation within different auditory areas, employing a wide range of stimulus categories.

The use of an fMRI adaptation paradigm has been suggested to be able to provide additional information about the representation of information in the brain (Grill-Spector et al. 2006), complementing the results obtained by classical fMRI experiments that contrast responses to different stimulus categories. The latter approach relies on spatially separable representations of different stimulus categories. However, it cannot always be assumed that the spatial separation between neural populations encoding different categories is in the range of the resolution of fMRI, in particular in higher order sensory areas. fMRI adaptation experiments are able to overcome these restrictions and allow to systematically test selectivity or invariance to various forms of transformations such as the effects of visual viewpoint changes in the ventral visual stream (James et al. 2002) or effects of size and position changes in object-related visual areas (Grill-Spector et al. 1999). Within the visual domain, fMRI adaptation effects have been replicated reliably in higher visual areas, whereas fMRI adaptation effects in early visual cortex are currently under debate. More specifically, although some researchers have reported orientation-specific fMRI adaptation in V1 (Tootell et al. 1998) and motion direction-specific fMRI adaptation in both lateral temporal–occipital and pericalcarine areas (Huettel and McCarthy 2004), other groups failed to identify orientation-specific fMRI adaptation in V1 (Boynton et al. 2003; Murray, Olman et al. 2006). Thus, it has been suggested that not only adaptation effects but also vascular mechanisms may contribute to the fMRI adaptation phenomenon. Recent studies (Sawamura et al. 2006) have employed an adaptation paradigm while acquiring single-cell recordings from macaque inferotemporal cortex for which fMRI adaptation has been observed previously (Sawamura et al. 2005). Although adaptation effects of single neurons were observed, (Sawamura et al. 2006) the tuning properties inferred from adaptation were sharper than the actual tuning of single neurons, suggesting that the link between adaptation and single-cell tuning curves is not straightforward. For the auditory domain, neural adaptation effects have been obtained for pure tones in cat primary auditory cortex (Ulanovsky et al. 2003), and combined human fMRI and MEG studies have provided evidence for stimulus-specific adaptation effects both for pure tones and noise (Jääskeläinen et al. 2004) and vowels (Ahveninen et al. 2006) in nonprimary auditory areas. Thus, employing adaptation paradigms might be a fruitful approach to investigate the representation of auditory information in the human brain. However, further research is needed to establish the link between different methodologies, such as single-cell recordings, human electrophysiology, and brain imaging.

Conclusions

In summary, our study provides evidence for the involvement of the bilateral STG in the processing of animal vocalizations. Moreover, our data suggest selective representation of animal vocalizations within the left STG as determined by an event-related fMRI adaptation paradigm. Adaptation effects were predominantly observed for putatively nonprimary auditory fields within the lateral STG rather than primary auditory cortex. The representation of animal vocalizations within this area is possibly mediated by the complex spectrotemporal dynamics of this stimulus class. These findings demonstrate that event-related fMRI adaptation is an appropriate method to test the selective representation of information in nonprimary
auditory fields. The investigation of cross-adaptation effects for different types of auditory features may lead to further insights into the organization principles of processing within auditory cortex.

Notes
This study was supported by the Bundesministerium für Bildung und Forschung (Brain Imaging Center Frankfurt, DLR 01GO0203). The authors would like to thank Tim Wallenhorst for assistance with data acquisition. Conflict of Interest None declared.

Address correspondence to Christian F. Altmann, Institute of Medical Psychology, Heinrich-Hoffmann-Straße 10, 60528 Frankfurt am Main, Germany. E-mail: C.Altmann@med.uni-frankfurt.de.

References


Huettel SA, McCarthy G. 2000. Evidence for a refractory period in the hemodynamic response to visual stimuli as measured by MRI. Neuroimage. 11:547-553.


