

Feature- and Object-based Attentional Modulation in the Human Auditory “Where” Pathway

Katrin Krumbholz^{1,2}, Simon B. Eickhoff¹, and Gereon R. Fink^{1,3}

Abstract

■ Attending to a visual stimulus feature, such as color or motion, enhances the processing of that feature in the visual cortex. Moreover, the processing of the attended object’s other, unattended, features is also enhanced. Here, we used functional magnetic resonance imaging to show that attentional modulation in the auditory system may also exhibit such feature- and object-specific effects. Specifically, we found that attending to auditory motion increases activity in nonprimary motion-sensitive areas of the auditory cortical “where” pathway. Moreover, activity in these motion-sensitive areas was also increased when attention was directed to a moving rather than a station-

ary sound object, even when motion was not the attended feature. An analysis of effective connectivity revealed that the motion-specific attentional modulation was brought about by an increase in connectivity between the primary auditory cortex and nonprimary motion-sensitive areas, which, in turn, may have been mediated by the paracingulate cortex in the frontal lobe. The current results indicate that auditory attention can select both objects and features. The finding of feature-based attentional modulation implies that attending to one feature of a sound object does not necessarily entail an exhaustive processing of the object’s unattended features. ■

INTRODUCTION

Attention is known to modulate activity in early sensory areas, enabling us to filter out the relevant from the irrelevant information that our senses are almost constantly bombarded with. In the visual modality, the cortical topography of attention-driven activity has been shown to mirror the topography of the purely sensory-driven activity elicited by the attended stimulus (for reviews, see Reynolds & Chelazzi, 2004; Pessoa, Kastner, & Ungerleider, 2003; Kastner & Ungerleider, 2000). Attending to a particular location, for instance, increases brain activity at the representation of the attended location in retinotopically organized areas (Brefczynski & DeYoe, 1999; Tootell et al., 1998). Visual attention can bias processing not only in favor of spatial locations but also in favor of nonspatial stimulus features, such as color or motion (Beauchamp, Cox, & DeYoe, 1997; Clark et al., 1997; Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990). O’Craven, Rosen, Kwong, Treisman, and Savoy (1997), for instance, have shown that attending to visual motion increases activity in the motion-sensitive medial temporal and medial superior temporal (MT/MST) cortex.

Although there is reliable evidence that attention to sounds can increase general activity in the auditory cortex (Lipschutz, Kolinsky, Damhaut, Wikler, & Goldman,

2002; Alho et al., 1999; Jäncke, Mirzazade, & Shah, 1999; Grady et al., 1997; O’Leary et al., 1997; Tzourio et al., 1997; Woodruff et al., 1996; Woldorff et al., 1993; Woldorff & Hillyard, 1991; Näätänen, 1990; Hillyard, Hink, Schwent, & Picton, 1973), feature-based attentional modulation has so far remained largely elusive in the auditory domain. Several studies have shown attention-driven activity in the auditory cortex to exhibit at least some degree of stimulus or task specificity but failed to provide definite proof for these effects to represent a modulation of feature-specific sensory-driven responses. For instance, the experimental design used in some of these studies does not allow for a direct comparison between attention-related and sensory-driven effects (von Kriegstein et al., 2003; Belin, McAdams, Smith, Savel, Thivard, & Samson, 1998), whereas other studies that did make this comparison have actually found the attentional and the sensory-driven activation to be localized in separate nonoverlapping areas (Petkov et al., 2004; Hall et al., 2000). In a series of studies, Woods and Alain (1993, 2001) and Woods, Alho, and Algazi (1994) investigated the effect of attending to feature conjunctions on the auditory-evoked potentials. They found that even sounds that possessed only one of the attended features can elicit an attention-related negative difference (Nd) potential (Hansen & Hillyard, 1980) and that Nd potentials to different features (frequency and location) exhibit different scalp topographies (see also Hansen & Hillyard, 1983). This would seem to suggest that the Nd potential to feature conjunctions reflects attentional modulation of

¹Research Center Jülich, ²MRC Institute of Hearing Research, ³University Hospital Aachen

feature processing. However, it has been shown that the Nd potential differs from the stimulus-driven response (N1) in its scalp topography and its adaptational properties, indicating that the Nd potential represents an additional response, the generators of which are not activated by unattended sounds, rather than a modulation of the stimulus-driven response (for a review, see Näätänen, 1990). The most convincing indication for feature-based effects in auditory attention comes from dichotic listening studies, which have shown that attending to one of two concurrent monaural sound streams, presented to different ears, increases activity in the hemisphere contralateral to the attended ear (Lipschutz et al., 2002; Alho et al., 1999; Tzourio et al., 1997). However, the contralaterality of monaural responses may be a mere epiphenomenon of the hemispheric asymmetry in the excitatory projections from subcortical levels and may, thus, not be representative of true feature processing in auditory cortex.

The difficulty of finding feature-based attentional modulation in the auditory cortex led Zatorre, Mondor, and Evans (1999) to suggest that auditory attention may operate at the level of object rather than feature representations. This hypothesis is supported by psychophysical data that show that performance in discriminating the pitch or location of an auditory stimulus is impaired by irrelevant changes in the respective other (unattended) dimension (Dyson & Quinlan, 2004; Mondor, Terrio, & Zatorre, 1998) and that there are costs associated with attending in two, rather than one, feature dimensions (Dyson & Quinlan, 2002). This indicates that attending to one feature of a sound object does not preclude the object's unattended features to also be processed to a certain degree. Although these results provide convincing evidence for object-based effects in auditory attention, their results do not reveal whether the unattended features of a sound object are processed to the same degree as the attended features, and thus, leave scope for the possibility that feature- and object-based attentional effects may coexist in the auditory system, as is the case in the visual system.

The aim of the current functional magnetic resonance imaging (fMRI) experiment was to revisit the search for feature-based attentional modulation in the auditory cortex using recent insights into the auditory functional anatomy. The experiment involved contrasting the responses to identical sound stimuli when attention was selectively directed to different stimulus features. To maximize the chances for detecting feature-based effects, we aimed to use stimulus features whose sensory processing would engage clearly separable cortical areas and an experimental design that would enable a direct comparison between attentional and sensory-evoked activity. To this end, we exploited the finding that object-related and spatial sound features are represented in different auditory cortical pathways, the anterior “what” and the posterior “where” pathways (Barrett & Hall,

2006; Arnott, Binns, Grady, & Alain, 2004; Hart, Palmer, & Hall, 2004; Warren & Griffiths, 2003; Alain, Arnott, Hevenor, Graham, & Grady, 2001; Maeder et al., 2001; Tian, Reser, Durham, Kustov, & Rauschecker, 2001; Rauschecker & Tian, 2000; Rauschecker, 1998). The to-be-attended features were associated with either the same or different sound objects in different conditions; this allowed us to also test for object-based attentional modulation.

METHODS

All stimuli were based on random noise, whose monaural temporal structure and/or interaural delay were manipulated to create stationary or moving sounds with or without pitch (tonal sounds and noises), which all had the same gross spectral structure. To create the tonal sounds, we changed the temporal fine structure of noise in a way that would regularize the time intervals between local peaks in waveform of the noise. The resulting sounds are referred to as regular-interval sounds (RISs); RISs have been used successfully in previous neuroimaging and electrophysiological studies on pitch perception (Krumbholz et al., 2003; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Griffiths et al., 1998). The pitch of the tonal sounds and the location of the moving sounds changed continuously over time. The experiment consisted of four test and three baseline conditions (Table 1). In the four test conditions, the stimuli possessed both pitch and motion. Participants were instructed to attend either to the pitch or to the motion and to indicate reversals in pitch-change or motion direction with a button press. In two of the test conditions (T1Pi and T1Mo), pitch and motion were

Table 1. Experimental Conditions

| <i>Condition</i> | <i>Features</i> | <i>Number of Sounds</i> | <i>Task</i> |
|------------------|---|-------------------------|-------------------------|
| <i>Baseline</i> | | | |
| BN | Stationary noise | One | None |
| BPi | Stationary <i>RIS</i> | One | Attend to <i>pitch</i> |
| BMo | Moving noise | One | Attend to motion |
| <i>Test</i> | | | |
| T1Pi | Moving <i>RIS</i> | One | Attend to <i>pitch</i> |
| T1Mo | Moving <i>RIS</i> | One | Attend to motion |
| T2Pi | Moving noise and stationary <i>RIS</i> | Two | Attend to <i>pitch</i> |
| T2Mo | Moving noise and stationary <i>RIS</i> | Two | Attend to motion |

Pitch and motion attributes are highlighted by *italic* and **boldface**, respectively.

attributes of a single sound object, a moving RIS. In the other two test conditions (T2Pi and T2Mo), pitch and motion were associated with different sound objects based on sequential perceptual segregation (Moore & Gockel, 2002), a stationary RIS and a moving noise, presented concurrently. In two of the three baseline conditions (BPi and BMo), the stimulus possessed pitch but no motion (stationary RIS; BPi) or motion but no pitch (moving noise; BMo) with the appropriate task as in the test conditions. In the remaining baseline condition (BN), the stimulus possessed neither pitch nor motion (stationary noise), and there was no task.

Stimuli and Experimental Protocol

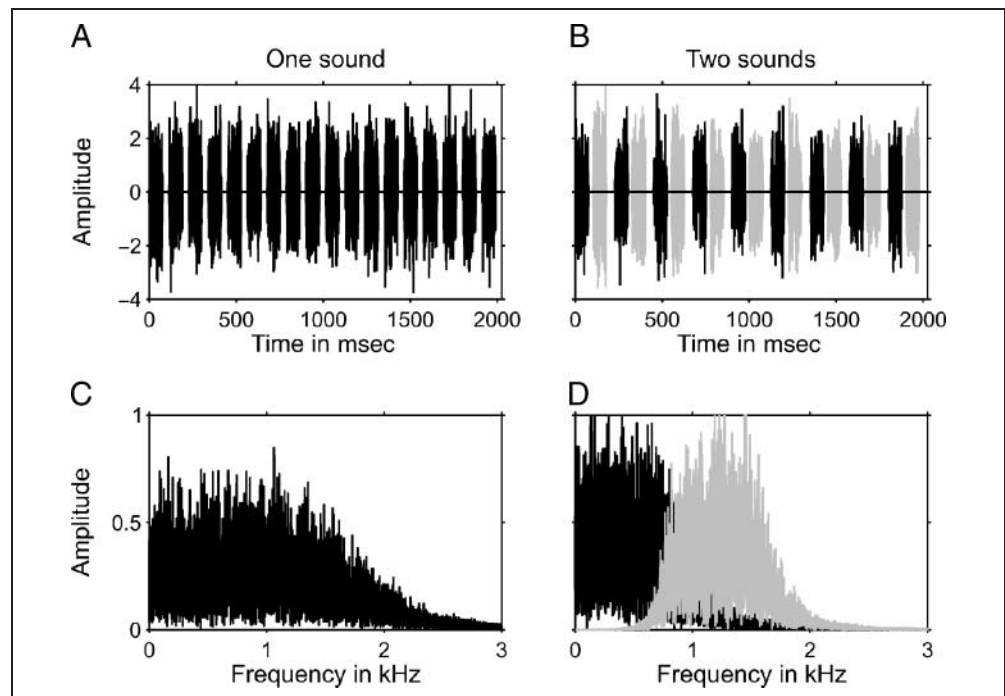
The experiment was divided into two runs, each consisting of 14 trials (two trials per condition per run). The order in which conditions were presented was counterbalanced across runs and across participants. Before each trial, a visual instruction was presented for 10 sec using a back-projection system. The instruction contained information about the stimuli that were going to be presented and the task that the participant was expected to perform. The stimuli were presented continuously for 40.5 sec. To avoid eye movements, participants were asked to fixate a cross at the midpoint of the visual axis, which was presented throughout the trials.

Stimuli were based on random noise, which was filtered to a single frequency band from 0 to 1.6 kHz in the single sound conditions or two contiguous bands from 0 to 0.8 and from 0.8 to 1.6 kHz in the two-sound conditions (Figure 1). The noise was generated afresh before each trial. The intensity of each of the two noise

bands in the two-sound conditions was equal to that of the single band in the one-sound conditions. To create the sensation of pitch, the temporal fine structure of the noise was regularized using a delay-and-add (to original) algorithm with four iterations and a gain of unity (Yost, Patterson, & Sheft, 1996). The delay was changed linearly over time to create the sensation of a pitch that either rose or fell continuously between 125 and 250 Hz for consecutive 2025-msec segments. Segments with rising and falling pitch were intermixed randomly in equal proportions, and in the attend-to-pitch conditions, participants were asked to respond with a button press (right index finger) whenever the direction of the pitch change had reversed. Stationary sounds were presented simultaneously to the two ears (zero interaural delay). The sensation of motion was created by changing the interaural delay linearly between -1000 and 1000 μsec for the same 2025-msec segments. By convention, a positive interaural delay means that the sound to the right ear is leading the one to the left. This created the sensation of a sound that moved from one ear to the other during 2025 msec and then either jumped back to the original ear or changed direction. Again, segments with leftward and rightward motion were randomly intermixed, and the participant was asked to indicate reversals in motion direction with a button press in the attend-to-motion conditions (same finger as in attend-to-pitch conditions). In the two-sound conditions, each type of sound (moving noise and stationary RIS) was equally often assigned to both frequency bands (0–0.8 and 0.8–1.6 kHz).

To maximize the stimulus-evoked response in auditory sensory areas and facilitate perceptual segregation

Figure 1. Temporal waveforms (A, B) and frequency spectra (C, D) of experimental stimuli. (Left) Moving RIS (used in the one-sound test conditions, T1Pi and T1Mo), whose temporal and spectral structure is also representative of the other stimuli that consisted of a single sound object (i.e., the stationary RIS and moving noise in the BPi and BMo conditions). (Right) Stimulus consisting of a moving noise (black lines) and a stationary RIS (gray lines) as used in the two-sound test conditions (T2Pi and T2Mo).



in the two-sound conditions, the sounds were pulsed on and off using 5-msec cosine-squared ramps. The sound duration (between the -6 -dB points) was always 75 msec. In the one-sound conditions, the sounds were presented once every 112.5 msec; in the two-sound conditions, each of the two sounds was presented once every 225 msec. Thus, the overall sound energy in the one- and two-sound conditions was equal. In the two-sound conditions, the two sounds were presented consecutively, with an interstimulus interval of 25 msec. Whether the sound in the lower (0–0.8 kHz) or higher frequency band (0.8–1.6 kHz) was delayed relative to the other sound was randomized across trials. Because of their nonoverlapping spectral pass bands and relatively fast presentation rate, the two sounds in the two-sound conditions were perceived as belonging to separate sound streams, originating from different sources (for a review on auditory sequential streaming, see Moore & Gockel, 2002).

The sounds were created digitally with a sampling rate of 12 kHz and 24-bit resolution using Matlab (The Mathwork, Natick, MA) and Tucker Davis Technologies System 3 (TDT, Alachua, FL) and presented to the participants through MR-compatible electrostatic transducers (Koss, Milwaukee, WI), built into professional ear defenders (Bilsom), which shielded the participant from the scanner noise (attenuation of at least 30 dB, even at low frequencies). The overall level of the stimuli was about 80 dB Sound Pressure Level (SPL). Care was taken that participants were able to hear all stimuli well enough to perform the task easily.

fMRI Data Acquisition

The noise associated with fMRI poses a problem for auditory experiments because it may mask the experimental stimuli acoustically and also produces a blood oxygen level-dependent (BOLD) response, which may overlap with the stimulus-related response. Sparse imaging, which uses a long image repetition time (TR), introducing periods of silence between successive images, is often used to avoid these problems. However, as we intended to perform an analysis of effective connectivity on the current data and as connectivity analyses are based on the variations in activity within experimental conditions (see Horwitz, 2003, McIntosh, 2000, Friston, 1994, and references therein), continuous imaging (short TR) rather than sparse imaging (long TR) was used in the current study. Although sparse imaging increases the signal-to-noise ratio in auditory fMRI data (e.g., Hall et al., 1999), each epoch contains only very few images (one or two), making it difficult to sample physiological variations in the hemodynamic response; this would be unfavorable for connectivity analyses.

A total of 568 BOLD contrast images were acquired for each participant using a Siemens (Erlangen, Germany) Vision 1.5-T whole-body scanner and gradient-echo-planar imaging (TR = 2.52 sec, echo time = 66 msec). The im-

ages were acquired continuously. Each image consisted of 25 four-millimeter slices with an interslice gap of 0.4 mm and an in-plane resolution of 3.125×3.125 mm², acquired in ascending order. A high-resolution structural image (3-D Magnetization Prepared Rapid Gradient Echo [MP-RAGE]) was acquired at the end of each measurement.

Data Analysis

Functional and structural data were analyzed with SPM99 (www.fil.ion.ucl.ac.uk/spm). The first four images were discarded to allow for magnetic saturation. The remaining images were realigned, slice time corrected, coregistered with the structural images, normalized, and smoothed using a Gaussian kernel with 10-mm full width at half maximum. The data of each participant were modeled within the framework of the general linear model. Each of the seven experimental conditions, as well as the instruction periods, was modeled as a box-car function convolved with the canonical hemodynamic response function (HRF). The data were high-pass-filtered (cutoff period = 350 sec) and precolored (by low-pass filtering with the HRF) to increase the hemodynamic variance components with a neuronal basis and account for serial correlations in the data (Friston et al., 1995). Contrast images for each participant were submitted to voxelwise one-sample *t* tests (one-tailed). The resulting random-effects *t* maps were thresholded at $t = 3.61$ ($p \leq .001$, uncorrected) and compared with the cytoarchitectonic areas on Heschl's gyrus using the SPM Anatomy toolbox (www.fz-juelich.de/ime/SPM_anatomy_toolbox; Eickhoff et al., 2005). The size of the signal change in the auditory cortex is often comparatively small, particularly when using continuous imaging. That is why many of the previous auditory fMRI studies used fixed—rather than random-effects analyses. Fixed-effects tests are based on the intraparticipant variance rather than the usually much larger interparticipant variance, and thus, usually yield a much more lenient threshold criterion than random-effects tests, even when probabilities are corrected for multiple comparisons. This is why we chose to use a random-effects analysis in the current study. It should also be pointed out that most of the relevant clusters in all of the relevant contrasts remained significant even after correction for multiple comparisons across the scanned volume (see Table 2).

To test for attentional modulation, individual β values for the experimental conditions were extracted from the most significant voxels in the contrasts testing for the sensory-driven effects of pitch and motion processing for the group data (56, -10 , and 2 and 60, -36 , and 12 mm, respectively, see Results and Table 2) and differences between the relevant conditions submitted to one-tailed *t* tests.

Attention-dependent contributions of the right primary auditory area TE1.0 and the motion-sensitive area in the region of the right temporo-parietal junction (TPJ) were assessed by testing for psychophysiological inter-

Table 2. Montreal Neurological Institute (MNI) Coordinates, *t* Values, Sizes (*k* in Number of Voxels), and Corrected *p* Values of Activation Clusters

| Contrast | Brain Region | Coordinates | | | <i>t</i> | <i>k</i> | <i>p</i> (Corrected) |
|---------------------|------------------|-------------|------------|-----------|--------------|-------------|----------------------|
| | | <i>x</i> | <i>y</i> | <i>z</i> | | | |
| Sensory pitch | Right STG | 56 | -10 | 2 | 8.22 | 1606 | .002 |
| | Left STG | -54 | -18 | 4 | 7.28 | 1196 | .008 |
| | Left PFC | -48 | 48 | -12 | 5.01 | 231 | .233 |
| | Right PFC | 54 | 18 | 14 | 4.67 | 112 | .367 |
| Sensory motion | Right TPJ | 60 | -36 | 12 | 6.73 | 535 | .048 |
| | Left STG | -46 | -30 | 8 | 5.56 | 108 | .265 |
| | Precuneus | 4 | -52 | 50 | 4.56 | 77 | .777 |
| | Left TPJ | -62 | -50 | 16 | 4.31 | 39 | .893 |
| Attention to motion | Right TPJ | 66 | -42 | 14 | 11.76 | 3192 | <.001 |
| | Precuneus | 4 | -50 | 56 | 9.42 | 1841 | .001 |
| | Right SMA/PMC | 26 | -4 | 62 | 7.27 | 2784 | .01 |
| | Right PFC | 28 | 48 | -16 | 7.19 | 94 | .012 |
| | Left TPJ | -64 | -26 | 38 | 6.50 | 1889 | .032 |
| | Left SMA/PMC | -32 | -8 | 48 | 5.59 | 787 | .131 |

PFC = prefrontal cortex; SMA = supplementary motor area; PMC = premotor cortex. Superior temporal clusters in the region of the auditory cortex and clusters in the parietal cortex are in **boldface**.

actions (PPIs) in two separate analyses. The rationale for the choice of these two areas was that area TE1.0 (i.e., the primary auditory cortex) can be assumed to be the physiological “input gate” to the cortical auditory system, whereas the right TPJ is a nonprimary motion-sensitive region, which also showed the most pronounced attentional effect. For the PPI analyses, the individual activation time series for these areas were obtained by extracting the first principal component of all voxel time series from a 5-mm sphere around the center of the anatomical probability map of primary area TE1.0 (51, -14, and 2 mm; Eickhoff et al., 2005) and the most significant voxel in the contrast testing for sensory motion processing for the group data (60, -36, and 12 mm, see Results and Table 2), respectively. The time series were high-pass-filtered (350 sec), mean-corrected, and multiplied with a vector representing attentional condition (1 for attend to motion, -1 for attend to pitch, and 0 elsewhere) to produce the PPI regressor. The attention vector was convolved with the HRF before multiplication. Next to the mean-corrected PPI regressor, the models also included the respective activation time series and the attention vector as regressors of no interest. To test for PPIs, *t* contrasts that were 1 for the PPI regressor and 0 elsewhere were computed, and the resulting individual contrast images were submitted to a random-effects group analysis using the same height threshold ($t = 3.61$) as used in the conventional analysis.

Participants

Nineteen right-handed (Oldfield test) volunteers (10 men and 9 women) between 20 and 35 years old with no history of audiological, psychiatric, or neurological disease participated in the experiment after having given written informed consent. The experimental procedures were approved by the local ethics committee.

RESULTS

Behavioral Data

To prepare the participants for the task, each participant was asked to perform the experiment behaviorally in a sound-treated booth (IAC) 1 or 2 days before the fMRI session. Hit minus false-alarm rates (Hit - FA) and response times (RTs) were submitted to a two-way repeated-measures analysis of variance with session (behavioral, fMRI) and condition (BPI, BMo, T1Pi, T1Mo, T2Pi, and T2Mo, see Table 1) as independent within-participant factors. Both analyses revealed a significant main effect of session, which was because of the Hit - FA rate being larger [$F(1, 18) = 6.329, p = .021$] and RT being smaller [$F(1, 18) = 4.404, p = .05$] in the fMRI compared with the behavioral session and thus probably reflects a training effect. The performance improvement in the fMRI compared with the behavioral session confirms that the continuous scanner noise during the fMRI session had little or no effect on the difficulty of the task, which is consistent

with participants' own reports. Moreover, the interaction between session and condition was insignificant for both Hit – FA rate [$F(5, 90) = 0.656, p = .658$] and RT [$F(5, 90) = 1.852, p = .111$], indicating that the scanner noise did not have different effects on different task conditions. Figure 2 shows the Hit – FA rates (A) and RTs (B) averaged across both sessions. The figure shows that there was also a highly significant main effect of condition on both Hit – FA rate [$F(5, 90) = 8.783, p < .001$] and RT [$F(5, 90) = 37.267, p < .001$]. Most participants reported finding the motion task more difficult than the pitch task. Although the Hit – FA rate was slightly larger for the motion than for the pitch conditions (see Figure 2A; significant differences are marked by stars), RT was overproportionately longer (Figure 2B).

In the attend-to-motion conditions, the sound was perceived as moving smoothly from one ear to the other during the 2025-msec cycles. It then “jumped” back to the original ear when there was no change in motion direction. At least part of the longer RTs in the attention-to-motion conditions may have been because of participants listening longer into the next cycle so as not to confuse this “jump” with a real change in motion direction.

Imaging Data

Sensory and Attentional Effects

To reveal activity associated with the sensory-driven processing of pitch or motion, in conditions where

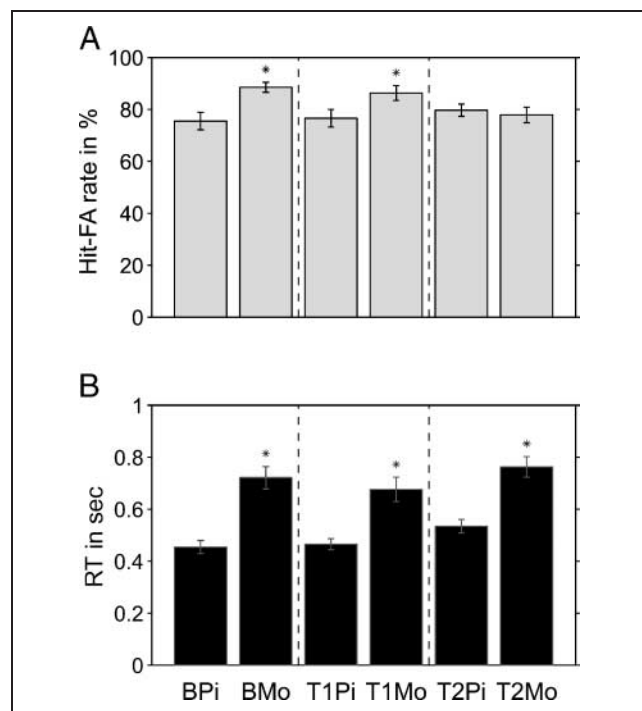


Figure 2. Hit – FA rates (A) and RTs (B) averaged across both the behavioral and fMRI sessions and plotted as a function of task condition. Error bars show the standard error of the mean across participants. Significant differences are marked with asterisks.

the respective feature was not attended to, each of the single-sound test conditions was compared with the baseline condition matching its respective attentional task. Thus, activity associated with the sensory processing of pitch was assessed by contrasting the moving RIS with attention-to-motion test condition with the moving-noise baseline condition (T1Mo–BMo); in both conditions, attention was directed to the motion, the only difference being the presence of an unattended pitch in the moving RIS. Similarly, activity associated with the sensory processing of motion was assessed by contrasting the moving RIS with attention-to-pitch test condition with the stationary-RIS baseline (T1Pi–BPi). As expected from previous results on object-related and spatial processing in the human (Barrett & Hall, 2006; Hart et al., 2004; Warren & Griffiths, 2003; Alain et al., 2001) and monkey auditory cortex (Tian et al., 2001; Rauschecker & Tian, 2000; Rauschecker, 1998), activity associated with the sensory processing of pitch (green in Figure 3A) was largely confined to regions anterior to those involved in the sensory processing of motion (red), the overlap between them (blue) being minimal. Sensory pitch processing produced an extended activation in anterior regions of the supratemporal auditory cortex bilaterally, comprising the anterolateral part of the planum temporale (PT), the lateral and central parts of Heschl’s gyrus, and the planum polare. In the right hemisphere, the activation extended downward and backward into the superior temporal sulcus (see sagittal slice in Figure 3A). Activity associated with the sensory processing of motion comprised the right and, to a lesser degree, the left TPJ, which refers to the region where the posterior part of the PT and the superior temporal gyrus (STG) adjoin the inferior parietal cortex, the anteromedial part of the PT on the left, and part of the right medial parietal cortex, henceforth referred to as the precuneus (PC) (see Figures 3A and 4B).

An analysis of the sizes of the BOLD effects for the different task conditions at the most significant voxels in the sensory pitch (56, –10, and 2 mm; Figure 3B) and motion contrasts (60, –36, and 12 mm; Figure 3C) revealed that activity in the motion-sensitive posterior areas was significantly modulated by attention. The response was larger in the attend-to-motion (gray and black bars with red stem in Figure 3C) than in the attend-to-pitch test conditions (gray and black bars with green stem). Importantly, this was true not only for the two-sound test conditions [Comparison 2 in Figure 3C, D; T2Mo > T2Pi: $t(18) = 6.19, p < .001$], but also for the one-sound test conditions [Comparison 1; T1Mo > T1Pi: $t(18) = 2.63, p = .008$]. However, although the attentional effect for the one-sound test conditions (Comparison 1) was statistically significant, it was only about half as large as the effect for the two-sound test conditions (Comparison 2). This difference was mainly because of the response to the one-sound attend-to-pitch test condition (gray bar with green stem in Figure 3C) being larger than the response to the two-sound attend-to-pitch condition [black

Figure 3. Sensory-evoked and attention-related activity.

(A) Activity associated with the sensory processing of pitch (green) and motion (red), projected onto an oblique axial (parallel to the Sylvian fissure) and a sagittal ($x = 60$ mm) slice of the Montreal Neurological Institute (MNI) single-subject template with probability maps of the primary auditory fields TE1.0 (dark gray), TE1.1 (light gray), and TE1.2 (white) (Morosan et al., 2001); the overlap between the pitch- and motion-specific activations is shown in blue. The yellow arrows mark the approximate positions of the most significant voxels in the sensory pitch and motion contrasts. (B, C) BOLD effect sizes for the different task conditions, including two of the baseline conditions (BPi and BMo; white bars) as well as the one-sound (T1Pi and T1Mo; gray bars) and two-sound test conditions (T2Pi and T2Mo; black bars), measured at the most significant voxels in the sensory pitch (B) and motion (C) contrasts (yellow arrows). Bars representing attend-to-pitch and attend-to-motion conditions are marked with green and red stems, respectively. (D) Contrast-weighted β values representing feature- (gray, red-hatched bars) and object-based attentional modulation (black, red-hatched bar) at the most significant voxel in the sensory motion contrast (Comparisons 1 to 3 in C). For comparison, the white bars with red and green hatching show the sizes of the sensory motion and pitch contrast, respectively (Comparison 4 in C and 5 in B).

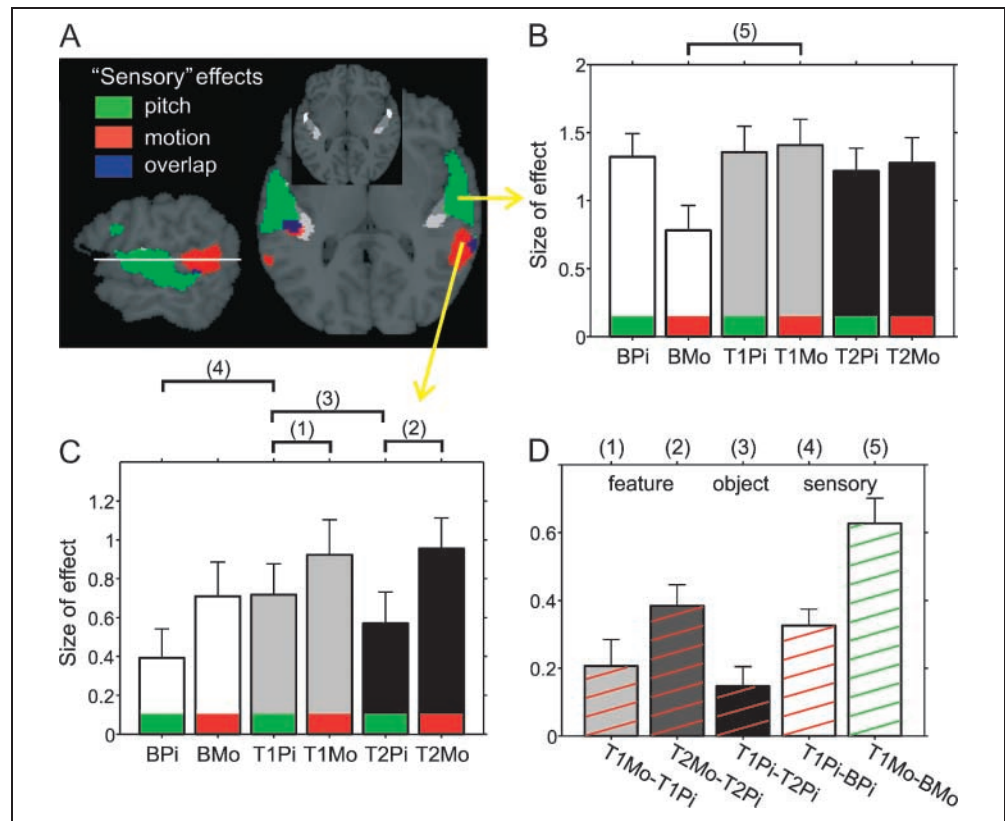
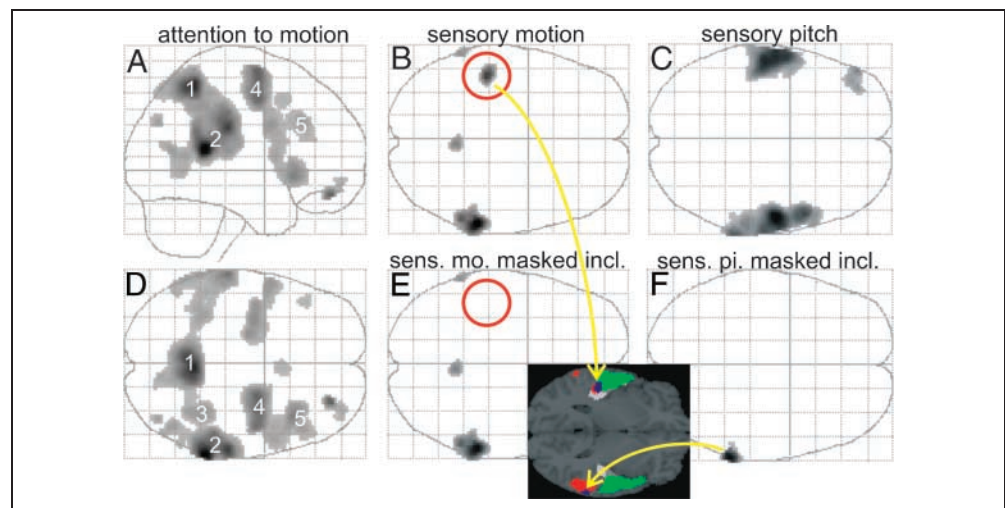


Figure 4. Scope and specificity of the attentional effect. (A, D) Sagittal and axial projections of the activation for the attention-to-motion contrast. (B, E) Sagittal projections of the sensory motion contrast and the sensory motion contrast masked inclusively by the attention-to-motion contrast. (C, F) Sagittal projections of the sensory pitch contrast and the sensory pitch contrast masked inclusively with the attention-to-motion contrast. The inset shows that the unmodulated part of the motion-sensitive area and the modulated part of the pitch-sensitive area were both part of the overlap between the pitch- and motion-sensitive areas (blue, compare Figure 3A).



bar with green stem in Figure 3C; see Comparison 3 in Figure 3D; $T1Pi > T2Pi$; $t(18) = 2.52$, $p = .011$]; the difference between the one- and two-sound attend-to-motion conditions (gray and black bars with red stem in Figure 3C) was insignificant.

Although the effect of sensory pitch processing (Comparison 5 in Figure 3B, D) was about twice as large as the effect of sensory motion processing (Comparison 4 in Figure 3C, D), activity in the sensory pitch area was similar irrespective of whether attention was directed to pitch (gray and black bars with green stem in Figure 3B) or to motion (gray and black bars with red stem), indicating that the pitch-sensitive areas did not show any effect of attending to pitch.

Scope and Specificity of the Attentional Effect

Contrasting the attend-to-motion with the attend-to-pitch test conditions ($T1Mo + T2Mo > T1Pi + T2Pi$) revealed all brain areas that were more strongly activated during the motion than during the pitch task (Figure 4A, D). This contrast will henceforth be referred to as the attention-to-motion contrast. Areas activated by this contrast included the medial parietal cortex bilaterally (1), the TPJ bilaterally, but stronger on the right (2), the right superior parietal cortex (3), supplementary motor and premotor areas bilaterally, but stronger on the right (4), as well as other prefrontal areas, mainly on the right (5). Most of these areas would be assumed to be related to processes of executive attentional control rather than the modulation of sensory processes.

Masking the attention-to-motion contrast (Figure 4A, D) inclusively with the sensory motion contrast (Figure 4B) showed that most of the motion-sensitive areas were also modulated by attention to motion (Figure 4E), apart from one small area in the anteromedial part of the left PT (encircled in red in Figure 4B, E). The unmodulated area was identical with the area of overlap between the sensory pitch and motion contrasts in the left hemisphere (compare blue clusters in inset and Figure 3A). This indicates that the modulatory effect of attending to motion encompassed all motion-sensitive areas that were not at the same time sensitive to pitch. Note that the inclusive masking of the attention-to-motion and sensory motion contrasts does not represent a statistical conjunction because the contrasts are not orthogonal. Rather, they share one element, $T1Pi$, to which they give opposite weights (negative in the attention-to-motion contrast, $T1Mo + T2Mo - T1Pi - T2Pi$; positive in the sensory motion contrast, $T1Pi - BPi$). This means that the effect of attention to motion is likely to be smaller the larger the effect of sensory motion processing. Because of this negative correlation, the inclusive mask may be expected to be more conservative than a true conjunction, and is thus very unlikely to overestimate the overlap between the two contrasts.

As the motion task was generally experienced as more difficult than the pitch task, there is a possibility that

the observed activation increase during the attend-to-motion conditions was simply an unspecific effect of task difficulty rather than a specific effect of attending to motion. However, an unspecific effect would be expected to not be limited to the motion-sensitive areas but rather affect all areas activated by the sounds, including the pitch-sensitive areas. To test whether the modulatory effect of attending to motion also affected the pitch-sensitive areas, the sensory pitch contrast (Figure 4C) was masked inclusively with the attention-to-motion contrast (Figure 4A, D). The resulting contrast (Figure 4F) yielded a small activation in the anterior part of the right temporo-parietal region. However, this activation was again part of the overlap between the pitch- and motion-sensitive areas and was thus also sensitive to motion (compare blue clusters in inset and Figure 3A). This indicates that the modulatory effect of attending to motion was highly specific to the motion-sensitive areas. Note that the inclusive mask may, in this case, be expected to exaggerate any overlap between the attention-to-motion and the sensory pitch contrasts (both contrasts give positive weight to $T1Mo$) compared with a true conjunction, thus reinforcing the claim that the effect of attending to motion was highly specific to the motion-sensitive areas.

In contrast, both the pitch and motion tasks did exhibit unspecific effects when compared with the passive baseline condition (BN): performing the motion task in the attend-to-motion baseline condition (BMo; stimuli possess no pitch) produced a significant activation increase ($p \leq .001$) relative to the passive baseline condition (BN) in 29% of the pitch-sensitive areas (921/3145 voxels; data not shown); similarly, performing the pitch task in the attend-to-pitch baseline condition (BPi; stimuli possess no motion) increased activation in 45% of the motion-sensitive areas relative to the passive baseline (343/759 voxels).

To further investigate the possible effect of task difficulty, two additional analyses were conducted using the Hit – FA rate for each trial or the Hit – FA rate normalized by the RTs [(Hit – FA)/RT] as parametric regressors and the attention conditions (attend to pitch and attend to motion) as regressors of no interest. According to Scheffe's post hoc tests, both the Hit – FA rate and the normalized Hit – FA rate exhibited significant differences between different attend-to-motion conditions [BMo vs. T2Mo ($p = .006$) and T1Mo vs. T2Mo ($p = .029$), respectively]. Nevertheless, neither of the two behavioral measures yielded any significant activation (at $p \leq .001$, uncorrected, as in original analysis) over and above the effects of the attention regressors (data not shown).

Attention-to-pitch Contrast

The reverse of the attention-to-motion contrast, testing for areas that were more strongly activated during the pitch than during the motion task, yielded no activation anywhere in the auditory system. The only activation

that this contrast produced was in the extrastriate visual cortex (Figure 5A, B). However, inspection of the BOLD effect sizes for the different task conditions at the most significant voxel in the attention-to-pitch contrast (34, -84, and -2 mm; Figure 5C) revealed that this apparent activation was actually because of a deactivation, with the deactivation being larger in the attend-to-motion (bars with red stem in Figure 5C) than the attend-to-pitch conditions (bars with green stem). This finding is consistent with previous results showing that difficult auditory tasks deactivate the visual cortex compared with easier tasks (Laurienti et al., 2002); in the current study, participants found the motion task more difficult than the pitch task (see behavioral data).

Changes in Effective Connectivity

Spatial hearing has been shown to involve a deep hierarchy of processing levels: whereas the actual perception of auditory space requires intact auditory cortical function (e.g., Clarke, Bellmann, Meuli, Assal, & Steck, 2000), the initial processing of auditory spatial cues, in particular, ITDs, begins as early as the brainstem (e.g., McAlpine, 2005; Grothe, 2003; Joris, Smith, & Yin, 1998). The current and previous results suggest that sensitivity to auditory motion mediated by time-varying interaural time delays (ITDs) emerges only toward the top of this processing hierarchy, that is, in nonprimary auditory areas (Krumbholz, Schönwiesner, Rübsamen, et al., 2005; Krumbholz, Schönwiesner, von Cramon, et al., 2005). This suggests that the processing of ITD-mediated motion is based on an analysis of successive instantaneous ITD estimates relayed from subcortical levels. This raises the hypothesis that the primary auditory cortex plays a key role in routing ITD-related information to the nonprimary motion-sensitive areas. To test this hypothesis, we assessed the effective connectivity of the right primary auditory cortex on medial Heschl's gyrus during the motion compared with the pitch task by means of PPIs (Friston et al., 1997). PPIs reveal where in the brain a given brain area exerts a stronger influence during one than during another experimental condition. Using linear regression, PPIs measure the "relatedness" of the activity in the given brain area with that in other areas; in this respect, PPIs are similar to correlational analyses (functional connectivity). However, by comparing the slope of the regression between different experimental conditions, PPIs enable to disambiguate artifactual correlations (e.g., produced by stimulus-evoked transients) from "real" correlations mediated by neuronal interactions (see Friston et al., 1997). Note that the definition of the volume of interest for the primary auditory cortex was based on anatomical rather than functional criteria (see Methods); the volume did not contain any voxels that were activated in the sensory motion or attention-to-motion contrasts. The primary auditory cortex was, however, more strongly activated by the task conditions

than by the no-task baseline condition (BN). Previous auditory fMRI studies found that the primary auditory cortex is generally activated by the presence of sound energy, irrespective of sound quality (e.g., Krumbholz, Schönwiesner, von Cramon, et al., 2005; Hall et al., 2002; Patterson et al., 2002). Testing for PPIs between attentional condition and activity in the right primary auditory cortex revealed an increased influence of this area on activity in the right PT, TPJ, and PC and, to a lesser degree, the left TPJ and right inferomedial thalamus during the attend-to-motion compared with the attend-to-pitch conditions (yellow in Figure 6A, B). Most of these regions were also activated in the sensory motion contrast (red) and were thus also modulated by attention to motion (see inclusive masking analysis, Figure 4). This suggests that the attentional modulation in the motion-sensitive areas was mediated by changes in their feed-forward input from the primary auditory cortex.

In a second analysis, PPIs for the motion-sensitive right TPJ were calculated to assess the connections of the nonprimary motion-sensitive areas with other, high-order brain areas. In the visual system, analyses of effective connectivity have demonstrated increases in connectivity not only from primary to nonprimary motion-sensitive visual areas (Friston & Büchel, 2000; Friston et al., 1997) but also from motion-sensitive areas to the parietal and frontal cortices (Penny, Stephan, Mechelli, & Friston, 2004; Büchel & Friston, 1997) during attend-to-motion conditions. A significant PPI between attentional condition and activity in the motion-sensitive right TPJ was found in the medial aspect of the superior frontal gyrus, above the anterior cingulate gyrus (cyan in Figure 6C, D); this area will henceforth be referred to as the paracingulate cortex. Interestingly, the paracingulate cortex was activated to a similar degree by all experimental conditions, including the no-task baseline condition (BN). In particular, activity in this area showed no main effect of attentional condition. This was shown by masking the contrasts between the relevant conditions inclusively with the paracingulate PPI activation; none of these masked contrasts yielded any significant activation (data not shown). This finding suggests that the paracingulate cortex may be involved in continuously monitoring attentional demand and mediating or controlling attentional modulation by increasing its coupling to task-relevant sensory areas.

The PPI for the motion-sensitive right TPJ and, albeit, to a much lesser degree, the right primary auditory cortex, yielded a significant deactivation in a similar area of the extrastriate visual cortex that also showed a deactivation in the attention-to-motion contrast (data not shown; compare Figure 5A, B).

DISCUSSION

Brain imaging and animal physiological data indicate that object-related and spatial sound features are represented

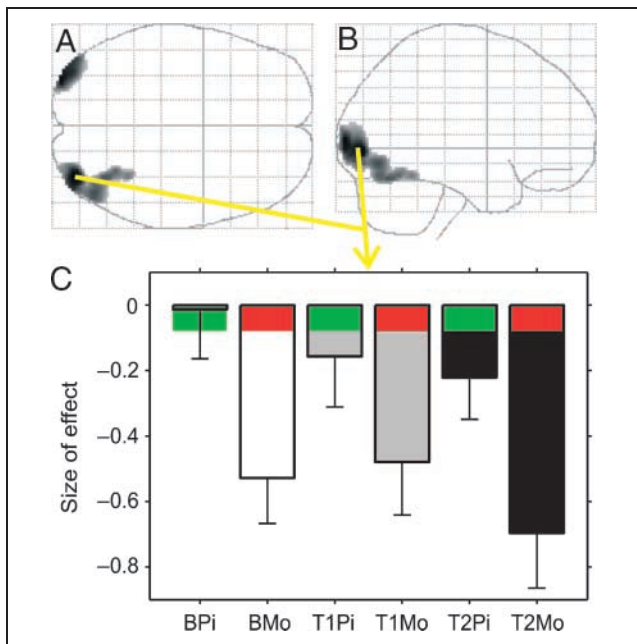


Figure 5. Attention-to-pitch contrast. (A, B) Axial and sagittal projections of the activation for the contrast between the attend-to-pitch and the attend-to-motion conditions. (C) BOLD effect sizes of the different task conditions at the most significant voxel in the attention-to-pitch contrast (see yellow arrows).

in different auditory cortical pathways, the anterior “what” and posterior “where” pathways (Barrett & Hall, 2006; Arnott et al., 2004; Hart et al., 2004; Warren & Griffiths, 2003; Alain et al., 2001; Maeder et al., 2001; Tian et al., 2001; Rauschecker & Tian, 2000; Rauschecker, 1998). The current data show that, in humans, voluntary attention to auditory motion modulates activity in those areas of the “where” pathway that are specifically involved in the

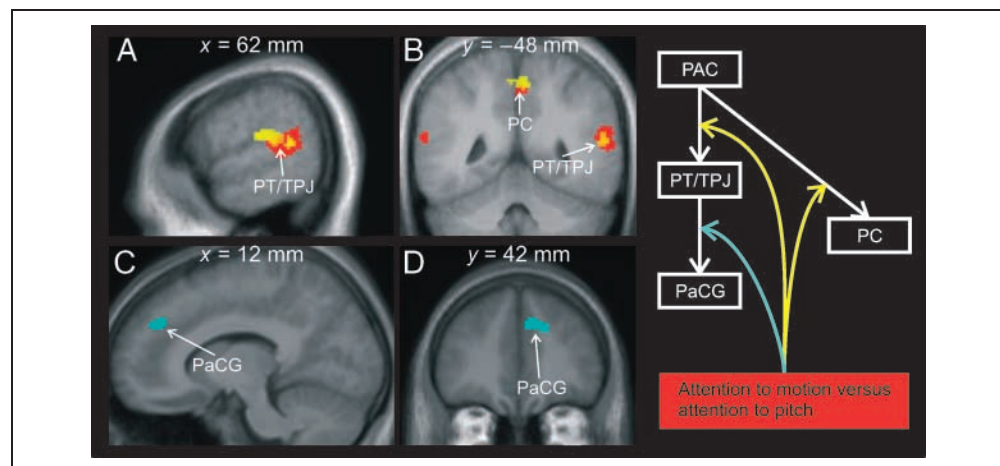
sensory processing of auditory motion cues. Attention to motion was contrasted with attention to pitch, an object-related feature that is known to be processed in areas of the anterior “what” pathway (Warren & Griffiths, 2003; Patterson et al., 2002). The attentional effect was highly specific to the motion-sensitive areas, and its size was comparable to, or even larger than, that of the effect of sensory motion processing in these areas (compare red-hatched bars with dark gray and white backgrounds in Figure 3D). An analysis of effective connectivity in the right hemisphere suggests that the attention-dependent increase in the response of the motion-sensitive areas was brought about by an increase in their functional coupling with the primary auditory cortex.

Feature-based and Object-based Attentional Modulation

Activity in the motion-sensitive areas was larger when attention was directed to motion than to pitch. This was true even for the one-sound test conditions where pitch and motion were attributes of the same sound object (moving RIS). The modulation was highly specific to the motion-sensitive areas, indicating that it was a feature-based rather than an unspecific effect. These results show that auditory attention can operate at the level of feature representations. Thus, as in the visual modality, attention to one feature of an auditory object does not necessarily entail the complete processing of the object’s other, unattended, features.

However, the observed attentional effect was about twice as large when pitch and motion were associated with different rather than the same sound object (two- vs. one-sound test conditions). This difference was mainly because of a difference between the one- and

Figure 6. Effective connectivity of primary and motion-sensitive auditory areas. Areas receiving an increased contribution from the right primary auditory cortex (PAC; area TE1.0) (A, B) and from the motion-sensitive right TPJ (posterior part of the PT and STG) (C, D) during the attend-to-motion compared with the attend-to-pitch conditions (compare schematic diagram on the right). (A, B) PPIs are highlighted in yellow. Most of the areas receiving attention-dependent contributions from the right primary area TE1.0 were also activated in the sensory motion contrast (red in A and B). PAC = primary auditory cortex (area TE1.0); PaCG = paracingulate gyrus.



two-sound attend-to-pitch test conditions rather than a difference between the attend-to-motion conditions. In the one-sound attend-to-pitch test condition, attention was directed to the pitch of a moving RIS. In this case, the unattended motion was a feature of an attended sound object. In the two-sound attend-to-pitch test condition, on the other hand, participants attended to the pitch of a stationary RIS, which was presented concurrently with a moving noise. In this case, the unattended motion was a feature of an unattended sound object. The difference in activation elicited by these two conditions in the motion-sensitive areas suggests that the unattended motion was more thoroughly processed when it was a feature of an attended rather than an unattended sound object. This finding is reminiscent of object-based accounts of attention in the visual modality (O'Craven, Downing, & Kanwisher, 1999, and references therein), suggesting that the difference between the attentional effects for the one- and two-sound test conditions was because of an object-based attentional modulation adding to the feature-based modulation in the two- but not the one-sound conditions. These results indicate that, as in the visual system, feature- and object-based attentional effects coexist in the auditory system.

Initially, we had expected to find a similar activation increase in the pitch-sensitive areas during the attend-to-pitch conditions as in the motion-sensitive areas during the attend-to-motion conditions. The absence of such pitch-related attentional modulation was most probably because of the pitch being more effective in capturing involuntary attention than the motion. Participants reported finding it more difficult to attend to the motion than the pitch, because the pitch tended to distract their attention more than the motion. The reason for this difference, which is also clearly reflected in the behavioral data, may be that pitch is behaviorally more salient than motion and auditory spatial cues in general. Behaviorally relevant sounds feature pitch almost without exception; pitch is the basis of prosody in speech and melody in music, and pitch is a salient cue for the simultaneous perceptual segregation of concurrent speech sounds, whereas interaural temporal cues, as used in the current experiment to create the perception of motion, are not (Culling & Summerfield, 1995). We expect that using a much weaker pitch in the pitch conditions (by reducing the number of iterations and/or the gain in the delay-and-add algorithm used to generate the RISs) would show similar attentional effects in the pitch-sensitive area as observed in the motion-sensitive area in the current study.

Brain Networks Underlying Attentional Modulation

Regions activated by the sensory processing of auditory motion cues were the TPJ bilaterally, the anteromedial PT on the left, and the right medial parietal cortex (PC). These regions are consistent with the results from

previous studies on auditory motion and spatial processing (Krumbholz, Schönwiesner, Rübsem, et al., 2005; Krumbholz, Schönwiesner, von Cramon, et al., 2005; Zimmer & Macaluso, 2005; Seifritz et al., 2002; Warren, Zielinski, Green, Rauschecker, & Griffiths, 2002; Zatorre, Bouffard, Ahad, & Belin, 2002; Griffiths, Green, Rees, & Rees, 2000; Baumgart, Gaschler-Markefski, Woldorff, Heinze, & Scheich, 1999; Bushara et al., 1999; Griffiths et al., 1998). Two of these regions, namely the right TPJ and PC, also showed the strongest modulation as a function of attention. The analysis of effective connectivity suggests that the attention-related changes in the activation of these two regions were brought about by changes in the strength of their connections to the primary auditory cortex, an area that was itself not specifically sensitive to motion or attention to motion. This finding is reminiscent of results from the visual system, where the connectivity between primary areas (V1/V2) and the motion-sensitive areas (MT/MST) has been shown to be modulated by attention to motion (Friston & Büchel, 2000; Friston et al., 1997). Another similarity is that, in the visual system, motion processing also seems to be subserved by a network of sensory (V1/V2, MT/MST), parietal, and frontal areas. Various approaches to estimate effective connectivity converge in suggesting that the visual motion processing network features a hierarchical system of reciprocal connections between the primary areas V1/V2 and areas MT/MST and between MT/MST and the parietal cortex (Penny et al., 2004; Friston, Harrison, & Penny, 2003; Büchel & Friston, 1997). The forward connection of each set of connections appears to be modulated by attention. Thus, both the visual and auditory motion networks seem to consist of similar components. However, in the visual network, the parietal cortex appears to be part of a hierarchical set of connections between these areas, whereas in the auditory network, there appears to be a connection between the parietal cortex and the primary sensory areas, and this connection forms a parallel branch to a more hierarchical set of connections between the primary and the nonprimary motion-sensitive areas and between the motion-sensitive areas and the frontal (paracingulate) cortex (Figure 6). The paracingulate cortex may be involved in continuously monitoring attentional demand, which would explain this area's approximately constant level of activation throughout all experimental conditions, and may be controlling attentional modulation by increasing its coupling to task-relevant sensory areas (compare Stephan et al., 2003).

Acknowledgments

We thank our colleagues from the magnetic resonance and the cognitive neurology groups at the Institute of Medicine for their support. This research was supported by the Helmholtz Association and a grant to G. R. Fink from the German Research Foundation (DFG-KFO 112).

Reprint requests should be sent to Katrin Krumbholz, Institute of Hearing Research, University Park Nottingham NG-7 2RD, UK, or via e-mail: katrin@ihr.mrc.ac.uk.

REFERENCES

- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 12301–12306.
- Alho, K., Medvedev, S. V., Pakhomov, S. V., Roudas, M. S., Tervaniemi, M., Reinikainen, K., et al. (1999). Selective tuning of the left and right auditory cortices during spatially directed attention. *Cognitive Brain Research*, *7*, 335–341.
- Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage*, *22*, 401–408.
- Barrett, D. J. K., & Hall, D. A. (2006). Response preferences for "what" and "where" in human non-primary auditory cortex. *Neuroimage*, *32*, 968–977.
- Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H. J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, *400*, 724–726.
- Beauchamp, M. S., Cox, R. W., & DeYoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *Journal of Neurophysiology*, *78*, 516–520.
- 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*, 6388–6394.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience*, *2*, 370–374.
- Büchel, C., & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex*, *7*, 768–778.
- Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., et al. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, *2*, 759–766.
- Clark, V. P., Parasuraman, R., Keil, K., Kulansky, R., Fannon, S., Maisog, J. M., et al. (1997). Selective attention to face identity and color studied with fMRI. *Human Brain Mapping*, *5*, 293–297.
- Clarke, S., Bellmann, A., Meuli, R. A., Assal, G., & Steck, A. J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: Evidence for distinct processing pathways. *Neuropsychologia*, *38*, 797–807.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *248*, 1556–1559.
- Culling, J. F., & Summerfield, Q. (1995). Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay. *Journal of the Acoustical Society of America*, *98*, 785–797.
- Dyson, B. J., & Quinlan, P. T. (2002). Within- and between-dimensional processing in the auditory modality. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1483–1498.
- Dyson, B. J., & Quinlan, P. T. (2004). Stimulus processing constraints in audition. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 1117–1131.
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, *25*, 1325–1335.
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, *2*, 56–78.
- Friston, K. J., & Büchel, C. (2000). Attentional modulation of effective connectivity from V2 to V5/MT in humans. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 7591–7596.
- Friston, K. J., Büchel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, *6*, 218–229.
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*, 1273–1302.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, *2*, 45–53.
- Grady, C. L., Van Meter, J. W., Maisog, J. M., Pietrini, P., Krasuski, J., & Rauschecker, J. P. (1997). Attention-related modulation of activity in primary and secondary auditory cortex. *NeuroReport*, *8*, 2511–2516.
- Griffiths, T. D., Green, G. G. R., Rees, A., & Rees, G. (2000). Human brain areas involved in the analysis of auditory movement. *Human Brain Mapping*, *9*, 72–80.
- Griffiths, T. D., Rees, G., Rees, A., Green, G. G. R., Witton, C., Rowe, D., et al. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience*, *1*, 74–79.
- Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, *4*, 540–550.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., et al. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping*, *7*, 213–223.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Summerfield, A. Q., Palmer, A. R., Elliott, M. R., et al. (2000). Modulation and task effects in auditory processing measured using fMRI. *Human Brain Mapping*, *10*, 107–119.
- Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *12*, 140–149.
- Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, *49*, 277–290.
- Hansen, J. C., & Hillyard, S. A. (1983). Selective attention to multidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 1–19.
- Hart, H. C., Palmer, A. R., & Hall, D. A. (2004). Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. *Human Brain Mapping*, *21*, 178–190.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Horwitz, B. (2003). The elusive concept of brain connectivity. *Neuroimage*, *19*, 466–470.
- Jäncke, L., Mirzazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: A functional magnetic resonance imaging study in human subjects. *Neuroscience Letters*, *266*, 125–128.
- Joris, P. X., Smith, P. H., & Yin, T. C. (1998). Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron*, *21*, 1235–1238.
- Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Reviews in Neuroscience*, *23*, 315–341.

- Krumbholz, K., Schönwiesner, M., Rübsamen, R., Zilles, K., Fink, G. R., & von Cramon, D. Y. (2005). Hierarchical processing of sound location and movement in the human brainstem and planum temporale. *European Journal of Neuroscience*, *21*, 230–238.
- Krumbholz, K., Schönwiesner, M., von Cramon, D. Y., Rübsamen, R., Shah, N. J., Zilles, K., et al. (2005). Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. *Cerebral Cortex*, *15*, 317–324.
- Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y. F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, *14*, 420–429.
- Lipschutz, B., Kolinsky, R., Damhaut, P., Wikler, D., & Goldman, S. (2002). Attention-dependent changes of activation and connectivity in dichotic listening. *Neuroimage*, *17*, 643–656.
- Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., et al. (2001). Distinct pathways involved in sound recognition and localization: A human fMRI study. *Neuroimage*, *14*, 802–816.
- McAlpine, D. (2005). Creating a sense of auditory space. *Journal of Physiology*, *566*, 21–28.
- McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Network*, *13*, 861–870.
- Mondor, T. A., Terrio, N. A., & Zatorre, R. J. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 66–79.
- Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica-Acustica*, *88*, 320–333.
- Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage*, *13*, 684–701.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral Brain Science*, *13*, 201–288.
- O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, *401*, 584–587.
- O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT-MST. *Neuron*, *18*, 591–598.
- O'Leary, D. S., Andreasen, N. C., Hurtig, R. R., Torres, I. J., Flashman, L. A., Kesler, M. L., et al. (1997). Auditory and visual attention assessed with PET. *Human Brain Mapping*, *5*, 422–436.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, *36*, 767–776.
- Penny, W. D., Stephan, K. E., Mechelli, A., & Friston, K. J. (2004). Modelling functional integration: A comparison of structural equation and dynamic causal models. *Neuroimage*, *23*, S264–S274.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *Journal of Neuroscience*, *23*, 3990–3998.
- Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience*, *7*, 658–663.
- Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinions in Neurobiology*, *8*, 516–521.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 11800–11806.
- Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience*, *27*, 611–647.
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Current Biology*, *12*, 2147–2151.
- Stephan, K. E., Marshall, J. C., Friston, K. J., Rowe, J. B., Ritzl, A., Zilles, K., et al. (2003). Lateralized cognitive processes and lateralized task control in the human brain. *Science*, *301*, 384–386.
- Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science*, *292*, 290–293.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron*, *21*, 1409–1422.
- Tzourio, N., El Massioui, F., Crivello, F., Joliot, M., Renault, B., & Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *Neuroimage*, *5*, 63–77.
- Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*, *23*, 5799–5804.
- Warren, J. D., Zielinski, B. A., Green, G. G. R., Rauschecker, J. P., & Griffiths, T. D. (2002). Perception of sound-source motion by the human brain. *Neuron*, *34*, 139–148.
- Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, U.S.A.*, *90*, 8722–8726.
- Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology*, *79*, 170–191.
- Woodruff, P. W. R., Benson, R. R., Bandettini, P. A., Kwong, K. K., Howard, R. J., Talavage, T., et al. (1996). Modulation of auditory and visual cortex by selective attention is modality-dependent. *NeuroReport*, *7*, 1909–1913.
- Woods, D. L., & Alain, C. (1993). Feature processing during high-rate auditory selective attention. *Perception and Psychophysics*, *53*, 391–402.
- Woods, D. L., & Alain, C. (2001). Conjoining three auditory features: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *13*, 492–509.
- Woods, D. L., Alho, K., & Algazi, A. (1994). Stages of auditory feature conjunction: An event-related brain potential study. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 81–94.
- Yost, W. A., Patterson, R. D., & Sheft, S. (1996). A time domain description for the pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America*, *99*, 1066–1078.
- Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is “where” in the human auditory cortex? *Nature Neuroscience*, *5*, 905–909.
- Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage*, *10*, 544–554.
- Zimmer, U., & Macaluso, E. (2005). High binaural coherence determines successful sound localization and increased activity in posterior auditory areas. *Neuron*, *47*, 893–905.