Examining the role of frequency specificity in the enhancement and suppression of human cortical activity by auditory selective attention

Aspasia E. Paltoglou

aspa.paltoglou@abdn.ac.uk

Christian J. Sumner

chris@ihr.mrc.ac.uk

Deborah A. Hall

deborah.hall@ntu.ac.uk

MRC Institute of Hearing Research

University Park

Nottingham

NG7 2RD

UK

Corresponding author: Deborah A. Hall

Correspondence address: Division of Psychology, School of Social Sciences, Burton Street, Nottingham, NG1 4BU. UK.

Telephone (with country and area code): +44 (0) 115 848 2377

Fax (with country and area code): +44 (0) 115 848 6829

Abstract

This study examined the neural basis of auditory selective attention using functional magnetic resonance imaging. The main hypothesis stated that attending to a particular sound frequency, would significantly enhance the neural response within those tonotopic regions of the auditory cortex sensitive to that frequency. To test this prediction, low- and high-frequency sound sequences were interleaved to produce two concurrent auditory streams. Six normally-hearing participants performed a task which required them to attend to one or the other stream, or they listened passively to the sounds while functional images were acquired using a high-resolution (1.5x1.5x2.5 mm) sequence. Two statistical comparisons identified the attention-specific and general effects of enhancement. The first controlled for task-related processes, while the second did not. Results demonstrated frequency-specific, attention-specific enhancement in the response to the attended frequency, but no response suppression for the unattended frequency. Instead, a general effect of suppression was found in several posterior sites, possibly related to resting-state processes. Furthermore, there was widespread general enhancement across auditory cortex when performing the task compared to passive listening. This enhancement did include frequency-sensitive regions, but was not restricted to them. In conclusion, our results show partial support for frequency-specific enhancement.

Keywords

Attentional modulation, frequency-dependent response region, task-related effects

Abbreviations

BOLD: blood oxygenated level dependent

CC: cingulate cortex

dB: decibel

IFG : inferior frontal gyrus

FDR: false discovery rate

FDRR: frequency-dependent response region

fMRI: functional Magnetic Resonance Imaging

FWHM: full width at half maximum

GLM: general lineal model

HG: Heschl's gyrus

Hz: Hertz

mm: millimeters

ms: microseconds

PT: planum temporale

s: seconds

SPL: sound pressure level

SPM: statistical parametric mapping

TE: echo time

TPOJ : temporo-parietal-occipital junction

Introduction

Selective attention is a fundamental aspect of cognitive activity that enables us to choose from the vast amount of information in the sensory world. The cognitive system has only a limited processing capacity and so behaviorally relevant information wins out for further processing while the rest is suppressed or ignored (Cherry, 1953). Studies of visual selective attention have established two important mechanisms that serve to optimize the perceptual contrast between the response to the object of interest and the background, namely enhancement (also referred to as gain) and suppression (Treue and Trujillo, 1999). In environments containing multiple objects defined by their conjunction of features, or single objects defined by multiple features, human fMRI studies have demonstrated feature-specific enhancement in regions that are sensitive to the attended features (Corbetta et al., 1990; 1991; Schoenfeld et al., 2007; Tootell et al., 1998) and feature-specific suppression in regions that are sensitive to the unattended features (Somers et al., 1999; Tootell et al., 1998).

The notions of feature-specific enhancement and suppression as the principle mechanisms of contrast enhancement are not restricted to the visual system. There is partial evidence that these mechanisms also exist in the auditory system, at least according to electrophysiological data recorded in awake animals when they are attending to sounds of a specific frequency. Sound frequency provides an excellent candidate for the study of feature-specific mechanisms of attention since it is the major organizing principle in the auditory system and neurons in primary auditory cortex are highly tuned to pure tones and narrowband sounds. Recordings of single-unit activity in awake, behaving ferrets have demonstrated frequency-specific changes in receptive field properties, consistent with attentional enhancement and suppression (Fritz et al., 2003; 2005; 2007a; 2007b).

In terms of human neuroimaging studies, evidence for attentional modulation in the auditory system is presently unclear. Using fMRI to quantify auditory cortical activity in nine listeners, Petkov et al. (2004) concluded that while responses in primary auditory cortex were tonotopically arranged, attentional enhancement occurred primarily in non-primary auditory cortex and was not frequency specific. This null result contradicts with other reports of significant feature-specific enhancement. For example, one fMRI study used a task that manipulated the focus of attention by instructing their nine listeners to discriminate either phonemes or spatial locations (Ahveninen et al., 2006). The results showed enhancement in non-primary auditory regions posterior to Heschl's gyrus (HG) when attending to the spatial feature, although there was no significant differential effect of attending to the nonspatial feature. For a group of 19 listeners, Krumbholz et al. (2007) again showed a significant increase in activity when attention was directed to spatial motion compared to pitch in posterior non-primary motion-sensitive areas, especially in the right temporo-parietal junction. A recent fMRI study has also confirmed the same feature-specific 'asymmetry' for spatial and nonspatial selective attention (Altmann et al., 2008). For a group of 12 listeners, the magnitude of feature-specific adaptation in motion-sensitive auditory cortical regions was influenced by whether attention was directed to location or to the spectrotemporal pattern, whereas adaptation in pattern-sensitive regions was not significantly modulated by the listening task.

At least two different experimental approaches have been used to reveal the effects of selective attention in auditory fMRI studies. The least-controlled method identifies any ‘general’ effect of attention by comparing an attention-demanding condition to a passive condition. In this type of experiment, the attention condition usually requires the participant to make a perceptual judgment, while the passive condition does not. Consequently, the resulting pattern of activity will include regions engaged in decision making and task execution, as well as selective attention. There are several auditory fMRI studies of this type (Ahveninen et al., 2006; Hall et al., 2000; Jäncke et al., 1999; Johnson and Zatorre, 2005). Better experimental control is achieved by effectively ‘subtracting away’ task-related differences by comparing two conditions that manipulate the focus of attention in different ways. We refer to this modulation as ‘attention-specific’. Auditory fMRI studies of this type include those in which listeners are required to direct their attention to either one feature of an auditory object or another (Ahveninen et al., 2006; Altmann et al., 2008; Krumbholz et al., 2007) or to either an auditory object or a visual object (Degerman et al., 2006; Johnson and Zatorre, 2005; Johnson and Zatorre, 2006; Petkov et al., 2004).

Despite the fact that the contrast which controls for other task-related activity is the preferred contrast for identifying attention-specific modulation in auditory cortex, we argue that it is important to also include the additional contrast comparing the attention conditions with passive listening. This argument is based on the context of a previous report. Specifically, Ahveninen et al. (2006) found, in the posterior auditory region not only an attention-specific effect of enhancement (comparing two active tasks), but also a general effect of enhancement (compared to passive listening). This finding would suggest a general elevation of the response during active listening that is perhaps not so indicative of an *attention-specific* effect. This type of interpretation could only be made by a study that includes both active and passive listening conditions.

The same form of logic can be applied to examine the effect of suppression. One can investigate the general effect of suppression by comparing a passive condition to one in which participants attend to the stimulus feature that is not preferentially coded by the brain region of interest. Alternatively, one can measure the attention-specific effects by contrasting two conditions that differentially manipulate the focus of attention (i.e. attending to the stimulus feature that is not preferentially coded by the brain region of interest versus attending to the ‘best’ stimulus feature. To the best of our knowledge, this aspect of attentional modulation has not been examined by auditory fMRI, but there are some good examples of general suppression from the visual literature (Somers et al., 1999; Tootell et al., 1998).

Although most demonstrations of feature-specific enhancement are for non-primary auditory cortical sites, not all studies report positive findings. For example in the study by Krumbholz et al. (2007) described above, there was no effect of attending to pitch in the pitch-sensitive areas. The null findings could be attributed to a range of methodological issues. In both studies by Petkov et al. (2004) and Krumbholz et al. (2007), the fMRI data were acquired at rather low resolution (voxel volume = 17.1 mm^3 and 39.1 mm^3, respectively) possibly chosen to maximize the signal-to-noise ratio on the 1.5 Tesla MR scanner. Furthermore in both studies, the continuous data acquisition (interscan interval = 2 s and 2.52 s respectively) generated an intense background noise that is known to particularly reduce the sensitivity to detect stimulus-driven activity in Heschl’s gyrus (Bandettini et al., 1998), the site of both frequency- and pitch-related activity. Interestingly, Krumbholz noted that the size of the attentional effect when attending to motion was greater (about twice as large) when the motion feature belonged to one of two auditory objects (presented as two concurrent streams) compared to when only one auditory object was presented that contained both motion and pitch. This effect can be explained in part by suggesting that attending to a single object with multiple features confers a ‘same-object’ advantage to the unattended features (O'Craven et al., 1999). In addition, the biased competition model predicts a greater effect of attentional enhancement when two competing objects are present at the same time, than when only one object is present (Desimone and Duncan, 1995; Luck et al., 1997). The size of any feature-specific enhancement may also be influenced by the difficulty in detecting or discriminating the attended feature (Atiani et al., 2009; Boudreau et al., 2006). In some cases, the null results may be attributed to the relative ease of the perceptual discrimination for the nonspatial feature (Altmann et al., 2008; Krumbholz et al., 2007).

The present experiment re-examines the evidence for frequency-specific attentional modulation using a carefully designed methodology. The task was made sufficiently difficult by manipulating the perceptual salience of the target so that listeners would be required to selectively attend to one stream and ignore the other. This was separately verified in a small group of listeners whom completed both selective and divided attention versions of the task. For the fMRI study, image acquisition was performed on a high-field 3 Tesla MR scanner using a high-resolution (voxel volume = 5 mm^3) scanning protocol in combination with sparse temporal sampling (interscan interval = 9 s, see Hall et al., 1999) to reduce the masking effects of the background noise. Frequency- and attention-dependent responses were mapped using the same two-step procedure reported by Tootell et al. (1998) which employed passive and active tasks, respectively to control for task-related differences. In the present study, although we do consider the general effect of enhancement, we prefer the strict definition of the ‘attention-specific’ effect which controls for the potential contribution of task-related differences to the effect of attention by considering the spatial relationship between the activity associated with the auditory stimulus features of interest and the effects of directing attention to one auditory feature or another. In other words, feature-specific attentional enhancement is defined as a relative increase in neural activity *in brain regions that are sensitive to the attended attributes of the stimulus compared to when attending to another attribute of the same stimulus that is not preferentially coded by that brain region.* In summary, the main hypothesis concerned the evidence for feature-specific attentional enhancement in the part of the tonotopic field best tuned to the frequency that was the focus of auditory selective attention. Conversely, we also tested the evidence for response suppression in the same part of the tonotopic field when attention was diverted towards the other auditory stream (i.e. attention was diverted ‘off best frequency’).

Materials and methods

Participants

Six right-handed, normally hearing (≤25 dB SPL at 0.25, 0.50, 1, 2, 3, 4 and 8 kHz) participants were recruited. The group comprised three males and three females, aged 19-29 years (mean = 24). The study was approved by the local Medical School ethics committee and all participants gave informed written consent. Four additional participants completed the behavioral screening, but their performance did not meet the inclusion criteria for the fMRI study.

Stimuli

The two concurrent auditory streams were each a 16-s stream containing a sequence of diotic narrowband 50-ms (5-ms onset and offset ramps) noise bursts. Both streams were perceived at a central spatial position, but could clearly be segregated in frequency. The low-frequency stream had a centre frequency of 0.25 kHz (bandwidth = 223-281 Hz) and the high-frequency stream had a centre frequency of 4 kHz (bandwidth = 3564-4490 Hz). Four-octave spacing was considered adequate for separating low- and high-frequency-dependent responses within tonotopic fields of the primary auditory cortex (Howard et al., 1996). As well as frequency, each stream was defined by a particular rhythm. The 'fast' rhythm comprised clusters of four bursts. Within each cluster, bursts were separated by 50-ms intervals of silence and between each cluster the interval was 250 ms. The 'slow' rhythm comprised single bursts separated by 550-ms intervals (Figure 1).

The two rhythms were counterbalanced across low- and high-frequency streams to generate two stimulus conditions. Henceforth, when describing the stimulus conditions, we use the term 'majority' to refer to the sound frequency occurring in the fast rhythm and 'minority' to refer to the sound frequency occurring in the slow rhythm. This choice of terms reflects the relative difference in the proportion of each sound frequency within the stimulus condition. Computing the frequency-dependent activity relied on the different relative proportions of low and high frequencies in the different stimuli.

*** insert Figure 1 about here ***

Task

A 2x3 factorial design crossed the two stimulus conditions with three listening conditions, with the further addition of a silent baseline condition. In the passive listening and baseline conditions, participants were asked to "just listen". During the fMRI experiment this was indicated by a fixation cross (‘X’). Note that there were no targets in the stimulus conditions that were used for passive listening as we did not wish to inadvertently draw attention to one stream or another.

Two further conditions directed attention either to the low-frequency stream or to the high-frequency stream (using the instructions ‘low pitch’ and ‘high pitch’, respectively). For the conditions that required directing attention to one of the streams, participants were required to perform a target-detection task in which the target was defined by an irregularity in the rhythm of the attended stream. The target in the majority stream occurred within the cluster and was defined by an inter-burst interval of 80 ms. The target in the minority stream was defined by an inter-burst interval of 615 ms. Piloting had shown that at these delays the target was detectable when attending to one stream, but was rather more difficult to detect when attending to the other stream. We reasoned that good performance would require sustained attention to the stream since targets could be detected only through a comparison between successive noise bursts and not the detection of isolated noise bursts. Each stream contained one or two targets and their occurrence was manipulated independently across the two streams. Sixteen different stimulus files were created for each experimental condition. Detection responses were recorded using an MR-compatible button box linked to the stimulus presentation computer for offline analysis.

Good task performance provides an objective marker for the ability to selectively attend to the target auditory stream. All participants completed two 30-minute screening sessions in a sound-attenuating booth and we selected only those participants who achieved a *d* prime score ≥ of 2 in all four ‘attend’ conditions, presented in two sessions in a pseudo-randomized order. To confirm that selectively attending to the target stream whilst ignoring the concurrent stream was more beneficial to performance than dividing attention simultaneously across the two streams, three participants (#4, #5 and #6) completed a second version of the task in the booth. In this version, participants were asked to detect targets whenever they occurred in the low- *and* high-frequency streams. Performance was directly compared to the selective attention version.

fMRI scanning was split into two 20-minute runs, each containing eight repeats of the six conditions and eight repeats of a silent ‘resting baseline’ condition. Sound files for each condition were presented in pairs, making each stimulus epoch 32 s in duration. Otherwise, conditions were presented in a randomized order. Stimuli were presented at 92 dB SPL using a bespoke set of MR-compatible electrostatic headphones. The instructions were visually presented using fiber-optic goggles (Silent Vision SV-7021, Avotec Inc.) mounted above the participant on the receiver head coil.

fMRI acquisition

Participants were scanned using a Philips 3 T whole body MR scanner (Achieva/Intera Release 1.2/11) equipped with a an 8-channel SENSE head coil. The fMRI session lasted about 1 ½ hours. A sagittal whole-brain anatomical scan was first acquired (1x1x1 mm, 256x256 matrix, 160 slices, TR=8.2 ms, TE=3.7 ms). This scan was used to position the subsequent functional scans (1.5x1.5x2.5 mm, 64x64 matrix, 18 slices, TR=9000 ms, TE=55 ms[1], flip angle = 90°) parallel to the Sylvian fissure with the central slice cutting through HG to maximize coverage of the supratemporal plane. To eliminate image artifact arising from the small field of view, saturation bands were positioned at the frontal and occipital poles roughly orthogonal to the imaging slice (Figures 2A and 2B). Each functional run contained 114 scans.

*** insert Figure 2 about here ***

fMRI analysis

Functional imaging data for each listener were first spatially transformed using standard procedures that are implemented in SPM2 software (www.fil.ion.ucl.ac.uk/spm). Individual time series were first motion corrected using rigid-body transformations (Cox and Jesmanowicz, 1999) to account for head movements both within and between the two runs using a central scan (i.e. the last scan of run 1) as a reference. Next the functional and anatomical scans were coregistered using a mutual information algorithm (Collignon et al., 1995; Studholme et al., 1998) to bring them into common alignment. To facilitate this step, a single 60-slice scan had been acquired in the same scanning session and using the same sequence parameters (except a shorter TE (i.e. 35 ms) to increase image brightness). This scan gave a greater field of view with which to better estimate the orientation of the functional data. For ease of localizing functional activity, image data were transformed into a standard brain space defined by International Consortium for Brain Mapping. In brief, the algorithm minimizes the sum of squares difference between the anatomical scan and a template scan by determining the optimum 12-parameter affine transformation and then by estimating the most appropriate nonlinear deformations based on a set of 3-dimensional discrete cosine basis functions (Ashburner and Friston, 1999). These same transformation parameters were applied to the functional scans. Normalized anatomical and functional scans preserved their original voxel resolution, although functional scans were smoothed using a 3 mm full-width at half-maximum (FWHM) Gaussian kernel to improve the signal-to-noise ratio.

Functional data for each listener were then modelled within the framework of the general linear model (GLM), as implemented in SPM2. The design of the model contained one regressor for each of the six experimental conditions, plus two additional regressors to account for the mean image intensity within each run. A high-pass filter with an 864-s (i.e. 0.001-Hz) cut-off was applied to the design to model low-frequency fluctuations in the time series. Comparisons between conditions estimated the statistical significance of each contrast with respect to within-subject variance.

*** insert Table 1 about here ***

Given the numerous statistical comparisons reported in the Results section, a detailed description of the contrasts is provided here in the same order as they are reported later. Initial analyses identified the low- and high-frequency dependent responses within tonotopically organised regions of the auditory cortex. To avoid any potential confound of attention, these analyses were applied to data acquired only during the passive listening state. The first of these contrasts (A>B, see Table 1) identified the low-frequency dependent response as those voxels showing a significant increase in activity for the sound condition containing a low-frequency majority stream compared to the sound condition containing a high-frequency majority stream. The second of these contrasts (B>A, see Table 1) identified the high-frequency dependent response in the converse manner.

Given that the main experimental hypothesis concerned a strict definition of the attention-specific effect, what we term the ‘attention-specific enhancement’ of auditory cortical activity used an appropriate control for the potential contribution of task-related differences. By contrasting two conditions in which the participants were performing a target-detection task in the context of ignoring the second concurrent auditory stream, we assumed that any task-related effects common to both conditions were subtracted away (Price et al., 1997). The first of these contrasts identified activity associated with directing attention to low frequencies. For this contrast, we selected the two low-frequency majority sound conditions in which the focus of attention was directed to either the low-frequency or the high-frequency stream (C>E, see Table 1). The second of these contrasts applied the same logic to identify activity associated with directing attention to high frequencies (F>D, see Table 1).

For comparison with some of the previous neuroimaging studies of selective attention (Ahveninen et al., 2006; Hall et al., 2000; Jäncke et al., 1999; Johnson and Zatorre, 2005), we also performed a subsidiary test of attention-related enhancement that was somewhat less stringent because it could not rule out the contribution of other task-related differences between active and passive listening states. This comparison is subsequently referred to as ‘general enhancement’. Again two contrasts were performed; one for attending to the low-frequency stream (C>A) and one for attending to the high-frequency stream (F>B, see Table 1). Note that each of these contrasts compared the condition in which attention was directed to the majority stream with the passive listening condition for the same stimulus.

The final test examined the evidence for response suppression. Theoretically speaking, one might expect that when attention is directed to the high-frequency stream, processing resources will be diverted away from those parts of the tonotopic field that are sensitive to low frequencies. This effect could be detected by contrasting the experimental condition in which attention is directed to the high-frequency minority stream (E) with the passive listening condition for the same stimulus (A) like so, E<A (see Table 1). Note that this interpretation of suppression is only valid for those auditory

cortical regions that have previously been defined as having a significant low-frequency dependent response. The second of these contrasts identified the effect of suppression in the high-frequency dependent response region (FDRR) caused by attending away from the high-frequency stream (D<B, see Table 1). Because these comparisons are not so conservative in their control for task differences, they are referred to as 'general' effects of suppression[2].

Due to the lack of a consistent voxel-wise mapping between anatomy and tonotopy across subjects, previous results reporting frequency-dependent activity have typically been done so separately for each participant (e.g., Talavage et al., 2000; 2004; Formisano et al., 2003). One way to examine the evidence for consistent patterns of frequency-related activity across participants is to tabulate the results in terms of standardized coordinates of peak activity (e.g. Talavage et al., 2000). Another method is to construct incidence maps by summing binary versions of the individual uncorrected statistical maps ($p<0.01$, see Hall and Plack, 2009). Incidence maps are informative because they describe the distribution of suprathreshold activity across participants in a manner that is visually accessible. Incidence maps were created for the contrasts specified in Table 1. The color scale of the incidence map denotes how many of the six participants showed significant activity at that coordinate. When the data from the six participants are combined in this way, the likelihood of false positive errors is greatly reduced. For example, the probability of activity in three out of six participants could be as small as 1.941×10^{-5}. Basing an interpretation only on those results that are consistent across a proportion of the participants in the group reduces the multiple comparisons problem. In reality, the true probability value is influenced by the degree of spatial correlation across voxels and this cannot be precisely determined. A second way in which the multiple comparisons problem was reduced was by restricting the number of voxels contributing to each incidence map. Maps showed only those voxels within anatomically specified borders (namely within a region encompassing the superior temporal gyrus and superior temporal sulcus, shown in Figure 2C).

Results

Target detection

A hit was considered to be a response made within two seconds after target offset. All other responses were taken as false alarms. Individual response bias was accounted for by calculating *d* prime measures of performance. To calculate correct rejections, each 16-s sound sequence was divided into eight 2-s bins. The *d* prime analysis for targets in the low-frequency stream was modified to account for the masking of targets that coincided with the burst of scanner noise by simply excluding this event from the analysis. Targets that were excluded were four of the 24 targets from the condition in which participants attended to the low-frequency targets in the low-frequency majority stimulus and three of the 24 targets from the condition in which participants attended to the low-frequency targets in the low-frequency minority stimulus. The success of participant training is shown in Figure 3A. Overall mean performance reached a *d* prime of 3.6 (stdev = 0.8). Performance significantly declined in the MR scanner ($Z=3.31$, $p<0.001$, Wilcoxon signed ranks test), although it still exceeded the criterion *d* prime of ≥ 2 (mean = 2.6, stdev = 0.8) (Figure 3B). After training, performance was not equivalent across all four conditions ($p<0.05$, Friedman test). Best performance was achieved when detecting high-frequency targets in the high-frequency majority stream and worst performance when detecting the same targets in the high-frequency minority stream. The same trend was observed during fMRI, although this pattern was not significant ($p>0.05$).

*** insert Figure 3 about here ***

All three participants who performed both selective and divided attention versions of the task showed a consistent performance cost for the latter version ($Z=2.20$, $p<0.05$, Wilcoxon signed ranks test) (Figure 3C). The mean *d* prime cost was 1.2 (stdev = 0.4).

Frequency-dependent responses

The incidence maps for the frequency-sensitive responses for low- and high-frequency sounds (A>B and B>A, respectively) across all six participants are shown in the top row of Figure 4. The most consistent frequency-dependent responses occurred in the area of primary auditory cortex on HG. Low frequencies primarily activated a strip across the posterior bank of HG. These peaks can be seen along the posterior (white) boundary of HG shown in Figure 4. The maximum incidence was 4 out of 6 on both left and right sides ($x = -56$, $y = -13$, $z = 5$ mm in left lateral HG and $x = 51$, $y = -16$, $z = 6$ mm in right central HG). These spatial co-ordinates were directly compared to the mean peaks of the frequency-dependent response regions (FDRR) reported by Talavage et al. (2000), after transformation into the same standard brainspace. A low-frequency site (named FDRR 1) was the closest peak and this measured 7 mm and 3 mm away, on left and right sides respectively. High frequencies mostly activated a focal site in the antero-medial part of HG, with a peak incidence of 3 out of 6 on the left ($x = -40$, $y = -22$, $z = 0$ mm) and 4 out of 6 on the right ($x= 39$, $y = -21$, $z = 8$ mm). Again, comparison with

Talavage et al. (2000) showed the closest peak to be a high-frequency site (named FDRR 2) and this measured 10 mm and 2 mm away, respectively.

*** insert Figure 4 about here ***

Attention-specific enhancement by selective attention

Next, we examined the effect of attention-specific enhancement (C>E and F>D). The incidence maps for these two contrasts are shown in the second row of Figure 4. Attention-specific enhancement was mainly restricted to primary auditory cortex. For the low frequencies, incidence peaks occurred in left lateral HG (x = -63, y = -19, z = 3 mm) and in right central HG (x = 48, y = -12, z = 2 mm). For the high frequencies, there was high incidence in the antero-medial part of left HG (x = -39, y = -21, z = 0 mm) and central part of right HG (x = 48, y = -13, z = 0 mm), extending to the central portion on the right side. For these peaks, the incidence ranged from 4 to 5 out of 6. No other regions of consistent attention-specific enhancement were observed across the brain.

We observed reasonably good commonality between the peaks of incidence and the tonotopic scheme supporting the hypothesis that attention operates by enhancing the frequency-dependent response. For example, the shortest Euclidean distance between the peak of FDRR 1 and the corresponding peak of the low-frequency attention-specific enhancement was 6 mm and between FDRR 2 and the corresponding high-frequency attention-specific enhancement it was a mere 1 mm.

Insert Table 2 and Figure 5 about here

Incidence maps provide at best an illustrative guide to the degree of commonality because they represent data pooled across all six participants. To quantify the extent of this commonality more rigorously within each individual, the number of suprathreshold (p<0.01) voxels was counted for each contrast of interest and the degree of overlap was computed across contrasts. To reduce the problem of multiple comparisons, this analysis was again restricted to voxels within the auditory cortical region of interest (shown in Figure 2C). The contrasts of interest were the frequency-dependent response and the two effects of selective attention (i.e attention-specific and general enhancement). Individual results are presented in Table 2 and mean results are reported in Table 2 and shown schematically in Figure 5. All participants demonstrated significant attention-related responses for low- and high- frequency contrasts. For the low frequency contrasts, of those voxels that exhibited an attention-specific enhancement, on average 15% (107/722) were also sensitive to low frequency sounds (ranging from 3.9% to 33.7% across participants). For high-frequency contrasts, this proportion was similar (15% or 117/771; range = 1.6 to 31.7%). Note that the maximum incidence for attention-specific enhancement did not reach 6 out of 6 (second row, Figure 4) but this does not necessarily establish a lack of effect in some participants. Instead, the incidence data reveal simply that the effect was not present in everyone at any one particular voxel.

In general conclusion, while our results confirm the presence of feature-specific attentional enhancement when attention is directed to the best frequency for that voxel, this is true for only a subset of voxels. In all six participants, by far the greater proportion of voxels showing attention-specific modulation occurred outside the frequency-sensitive regions, thus failing to support the strict definition of the attention-specific, frequency-specific effect of selective attention.

General enhancement by selective attention

Next we considered the general effect of enhancement when attending to the low- and high-frequency streams (C>A and F>B, respectively). The incidence maps for these two contrasts are shown in the third row of Figure 4. In both cases, there was patchy but widespread activity involving both HG and planum temporale (PT). For the low frequencies, numerous incidence peaks occurred around central and lateral portions of the posterior bank of HG, each with an incidence of between 3 and 4 out of 6. For the high frequencies, three sites of consistent activity were noted. One in the antero-medial part of HG bilaterally (on the left, the peak of incidence was 6 out 6, x = -39, y = -24, z = 0 mm, and on the right it was 3, x = 42, y = -20, z = 3 mm); one in the left postero-lateral part of HG with a peak of 4 (x = -57, y = -14, z = 5 mm), and one in the right postero-medial part of HG with a peak of 5 out of 6 (x = 54, y = -27, z = 10 mm). The limited choice of axial slices for display means that not all of the above peaks are clearly visible in Figure 4. However, we note that FDRR 1 generally demonstrated general attentional enhancement, albeit it with reduced incidence (2 out of 6). However, in the high-frequency site (i.e. FDRR 2), all six participants showed a general response enhancement on the left.

Unlike the attention-specific enhancement, general enhancement occurred predominantly in regions beyond primary auditory cortex. Notable effects occurred found in PT, consistent in at least 3 out of 6 participants. Activity here appeared not to be frequency-specific as it occurred for both low and high-frequency contrasts (e.g. in left PT, x = -48, y = -42, z = 25 mm and in right PT, x = 66, y = -36 z = 25 mm).

Insert Figure 6 about here

Preliminary analyses suggested some additional general enhancement in inferior frontal gyrus and insula bilaterally. For those unfamiliar with human neuroanatomy, the approximate position of these areas is illustrated in Figure 6. Given that these prefrontal regions were outside our anatomically constrained hypotheses, statistical significance was assessed with caution. Specifically, a correction for false discovery rate (FDR) (Genovese et al., 2002) was applied to the whole brain ($p<0.05$) for each individual low- and high-frequency contrast. For inferior frontal gyrus (IFG), when new incidence maps were created using the corrected threshold, enhancement for low and high frequencies was most consistent across participants in the left hemisphere. The maximum overlap was 4 out of 6 and the coordinates of maximum incidence are reported in Table 3. Both peaks fell within Broca's area (BA 44) (Amunts et al., 1999). For insula cortex, activity was reliably present in both hemispheres for low and high-frequency contrasts (see Table 3).

General effect of suppression

The bottom panel in Figure 4 displays the incidence maps for the two general effects of suppression (E<A and D<B) in auditory cortex. These maps indicate no reliable general suppression within auditory cortex.

Insert Figure 7 about here

Once again however, preliminary analyses suggested that effects of general suppression were present elsewhere in the brain. To examine this further, incidence maps were created using individual corrected T contrasts (false discovery rate, $p<0.05$). Non-auditory brain regions showing a reliable effect of suppression encompassed the posterior part of temporo-parietal-occipital junction (TPOJ) bilaterally and the precuneus and posterior cingulate cortex on the midline (Table 3 and Figure 7). Just like general enhancement, general suppression was frequency independent. Reliability of these effects was demonstrated by the peak of maximum incidence which was typically 4 out of the 6 participants.

Within these non-auditory regions, the pattern of activity was examined by quantifying the contribution of each stimulus condition (i.e. the effect size) to the overall variability in the fMRI signal. Effect size (β) is estimated for every voxel as part of the GLM and so it was possible to extract these values for all peak voxels reported in Table 3, separately for each participant. Summarizing these data across participants and across hemispheres, a pattern emerged that was consistent for both low- and high-frequency contrasts in all three regions of interest. To illustrate this pattern, the data for the low-frequency contrast are displayed in Figure 7. The 95% confidence intervals around the mean enable the reader to visually interpret which conditions statistically differ. We observed that the effect of suppression arose from a significant inhibition of activity (negative effect size) during the attention-demanding conditions, relative to passive listening. The functional interpretation of this result will be discussed later.

Discussion

The results from this sample of six participants support a mechanism of general auditory enhancement when the listening task required an 'active' perceptual judgement. In contrast to our original hypothesis, the most consistent result was that of a spatially extensive increase in activation that was frequency *in*dependent and involved both primary and posterior nonprimary auditory cortical regions. A proportional representation of the number of voxels that showed the different effects are summarized schematically in Figure 5. By far the greatest proportion of voxels in auditory cortex to show a significant attention-specific modulation in their response were those *not* significantly sensitive to sound frequency. Nevertheless, our results also lend partial support for an additional mechanism of frequency-specific enhancement in primary auditory cortex during auditory selective attention because there was reasonably good correspondence between the peaks of frequency-dependent activity and attention-specific modulation. Moreover, the lack of any significant task-related suppression is consistent with the view that selective auditory attention is dominated by general and enhanced sound processing activity.

Task-specific enhancement in primary auditory cortex

The findings within the proximity of human primary auditory cortex are generally consistent with a wide range of electrophysiological data recorded in primary auditory neurons of awake, behaving animals. Many such studies demonstrate that active engagement in a listening task increases the reliability of auditory responses (Benson et al., 1981; Pfingst et al., 1977). A large number of studies have also reported positive effects in auditory cortex when manipulating attention (Benson and Hienz, 1978; Fritz et al., 2003; Fritz et al., 2005; Hocherman et al., 1976; Hubel et al., 1959). In particular, a number of shock-avoidance, conditioning studies have shown a very rapid plasticity of frequency tuning in favor of the conditioned tone frequency (Bakin et al., 1990; Diamond and Weinberger, 1986; Diamond et al., 1989; Edeline and Weinberger, 1993; Edeline et al., 1993; Fritz et al., 2003; 2005; Ohl and Scheich, 1996). For example, Fritz et al. (2003) demonstrated frequency-specific enhancement of the response during a number of attention-demanding tasks using tone stimuli. Although these authors used a very different stimulus, these electrophysiological results give support to the existence of frequency-specific attentional enhancement in primary auditory cortex. Moreover, they do not predict specific enhancement of neurons tuned to target frequencies, but rather a recruitment of neurons not precisely tuned to the stimulus. Speculatively, this might be contributing to the lack of direct correspondence between sites of maximal frequency sensitivity and of attentional enhancement.

We acknowledge that comparisons between electrophysiological and fMRI studies must be treated with caution. While the above animal studies reported task-related effects in terms of enhanced action potentials, the present experiment reported task-related changes in blood-oxygen-level-dependent (BOLD) signal. The BOLD response is not necessarily correlated with the actual neural firing output of that region. In fact, it has been shown to be more typically associated with sub-threshold input to, and local processing within, a region because both of these processes place demands on energy metabolism (Logothetis and Wandell, 2004). Thus, single-unit recordings of action potentials and BOLD fMRI could measure different aspects of neural activity.

A number of electrophysiological studies in animals have also examined cortical responses to sequences of alternating tones (e.g. Bee and Klump 2004, Fishman et al., 2001, 2004; Micheyl et al. 2005). The results support the notion that concurrent auditory streams, separated by large frequency differences, are represented in segregated populations of neurons. Human neuroimaging studies have also suggested that the amount of auditory cortical activation differs according to whether the sound sequence is perceived as one or two streams (e.g., Gutschalk et al., 2005; Snyder et al., 2006; Wilson et al., 2007). For example, when the repeating sequence segregated into two streams, fMRI activation was more sustained throughout the sequence duration and was larger in magnitude and extent than when the repeating sequence was heard as a single stream (Wilson et al., 2007). There is little basis for comparing these results with those from the present experiment because they were specifically concerned with the neural basis of the streaming phenomena, while we simply used the streaming paradigm as a device for manipulating the focus of frequency-specific attention.

Although our data lend support to the hypothesis that attending to sound frequencies can modify response properties in human auditory cortex, it is not possible to say at this point whether or not such activity reflects enhancement to attended auditory features or objects. While several studies (e.g. Krumbholz et al. 2007; Fritz et al. 2005) do appear to support the hypothesis that selective attention can operate at the level of object-based representations, evidence that uniquely supports the notion of object-based auditory selective attention is still lacking. The present data would be consistent with attributing the observed changes with task to a feature-based attentional mechanism, especially for those voxels in the vicinity of primary auditory cortex where neurons are sharply tuned to frequency.

No evidence for task-specific suppression in primary auditory cortex

Electrophysiological studies have suggested that two-tone discrimination not only to increases firing rate of frequency-specific responses in primary auditory cortex, but also *decreases* firing rate (suppression) in that spectral region of the STRF corresponding to the non-target reference tone (Fritz et al., 2005). Our fMRI results are somewhat difficult to reconcile with these findings because we did not find any significant or consistent suppression in auditory cortex. It is interesting to note that Fritz et al. (2007a) noted a qualitative difference between enhancement and suppression. While enhancement persisted over a relatively long time after the task was completed (up to 45 minutes), suppression faded rather quickly.

The effect of sound level on fMRI mapping of frequency-dependent responses

In the present study, stimuli were presented at rather high sound levels and so it is worthwhile considering how this might influence the sensitivity to localising frequency-specific effects within tonotopically organized areas of auditory cortex. Loud tones are well known to stimulate additional neurons tuned to adjacent frequencies. On the low-frequency side in particular, tuning curves become progressively less sharply tuned with increasing sound level. Such an upward spread of excitation might be particularly expected to result in high-frequency regions responding to the low-frequency stimulus. While this certainly occurs, it does not unduly compromise our conclusions regarding the peak locations of low- and high-frequency dependent responses. First, the frequency separation between the low- and high-tones was about 4 octaves and so, even for sounds as loud as 92 dB SPL, the high and low-sensitive regions are unlikely to overlap considerably. Although the incidence maps did show a greater extent of activity for the low-frequency regions than the high frequency regions, nevertheless there was general separation between the two. Second, we note that in the present fMRI study ‘activity’ was defined as a differential (frequency-related) response between the low and high tones, not as a contrast between each tone condition and a silent baseline as had been done previously in our group (see Hart et al., 2002; 2003). A differential contrast is more sensitive to defining the peak of the frequency-related response and less sensitive to the tails of the response.

The role of nonprimary auditory cortex and prefrontal cortex in attention-demanding listening

Our data indicate that non-primary auditory cortex (especially PT) was engaged by the attention-demanding listening task. Regions within PT were significantly activated by general attention and not by the attention-specific contrast. This pattern of results is consistent with a role for PT in auditory selective attention, since one could argue that such attention is not required during passive listening but is required for all target-detection conditions. Griffiths and Warren (2002) proposed that PT might play a key role in auditory scene analysis, and although they did not rule out its potential contribution to an auditory attentional network, they claimed that its computational role did not necessarily depend on attention. In the present study, the target-detection task required stream segregation which is one form of auditory scene analysis. Therefore, while our data are certainly not inconsistent with this model, they cannot clearly distinguish whether the role of PT here is primarily one of stream segregation or whether it forms part of a general auditory attentional network.

Inferior frontal gyrus in the left hemisphere showed a similar pattern in its response, namely a reasonably consistent effect of general enhancement but no reliable attention-specific enhancement across the group. Given the status of prefrontal cortex as a hub for higher-order cognitive processes, we argue that it is unlikely to be the site of those neural computations required to parse different auditory objects within the scene. Instead, inferior frontal gyrus is one of a number of prefrontal regions that is consistently found to be recruited for solving a broad range of different cognitive problems, including aspects of perceptual discrimination, response selection, executive control, working memory, and problem solving (Duncan and Owen, 2000). It could be argued that all of these processes distinguish the active from the passive listening conditions. Likewise, activity in bilateral insula cortex was also enhanced during active compared with passive listening. While one of the main functions of the insula concerns the generation of subjective emotional experiences for sensory stimulation and for visceral states, this region has also been implicated in a range of attention-demanding tasks (Altmann et al., 2008; Sridharan et al., 2008). It has been suggested that inferior frontal gyrus and insula play a co-ordinated role in providing an enhanced level of attention directed towards multimodal stimuli during difficult tasks (Downar et al., 2001).

A putative cognitive role for TPOJ, precuneus and posterior cingulate cortex

Despite a lack of suppression in the auditory cortex, there was a reliable effect of suppression in a number of other regions arising from a negative response in all attention-demanding conditions. The functional significance of a negative response has been widely debated. There are a number of general proposals. One is that the negative response has a vascular origin in which passive decreases in arterial cerebral blood flow stem from a spatial redistribution of blood flow and a reduction in blood flow to nonstimulated brain areas (Harel et al., 2002). We suggest that, in our case the suppression is unlikely to reflect ‘blood steal’, since TPOJ precuneus and posterior cingulate cortex are all situated far from the auditory cortex and have no unique neurovascular links (Duvernoy, 1999).

Another explanation is that the negative response reflects a genuine local reduction in neural activity relative to its intrinsic spontaneous level (Shmuel et al., 2006). Although our indirect fMRI measure of brain activity cannot ascertain this for certain we favor this interpretation. We propose that the decreases might reflect the operation of a ‘default’ brain system that is highly active when the participants is at rest and deactivated when they are engaged in a task (Buckner and Vincent, 2007; Gusnard and Raichle, 2001). The default network involves precuneus and posterior cingulate cortex, and also prefrontal and lateral parietal regions. In a meta-analysis of nine neuroimaging studies of visual processing, Shulman et al. (1997) reported that this default-mode network consistently included TPOJ. Our results extend these findings to suggest that the default-mode network is modality independent. Note that the other target region in prefrontal cortex was not within the volume of the functional scanning.

Footnotes

[1] At the time of scanning, this value was considered to be the optimal echo time for auditory cortical grey matter using the chosen scanning sequence. Optimal TE is equivalent to the T2* relaxation rate and can be computed from a plot of MR signal intensity as a function of TE (T2* = -1/(gradient)). In practice, an error in our original calculations of the values generating the linear regression increased our estimate of the optimal TE. In itself, this does not invalidate our results since it was applied to all participants and to all stimulus conditions, but it perhaps does somewhat reduce the overall BOLD sensitivity.

[2] Given the lack of any significant effects of general suppression in auditory cortex, we did not perform the more conservative analysis that would have controlled for task differences between active and passive listening states.

References

Altmann, C.F., Henning, M., Döring, M.K, Kaiser, J. 2008. Effects of feature-selective attention on auditory pattern and location processing. NeuroImage 41, 69–79.

Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.H., Sams, M., Shinn-Cunningham, B.G., Witzel, T., Belliveau, J.W. 2006. Task-modulated "what" and "where" pathway in human auditory cortex. Proceedings of the National Academy of Sciences of the United States of America 103, 14608-14613.

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H.B.M., Zilles, K. 1999. Broca's region revisited: Cytoarchitecture and intersubject variability. The Journal of Comparative Neurology 412, 319-341.

Ashburner, J., Friston, K.J. 1999. Nonlinear spatial normalization using basis functions. Human Brain Mapping 7, 254-266.

Atiani, S., Elhilali, M., David, S.V., Fritz, J.B., Shamma, S.A. 2009. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields 61, 467-480.

Bakin, J.S., Weinberger, N.M. 1990. Classical conditioning induces CS- specific receptive field plasticity in the auditory cortex of the guinea pig. Brain Research 536, 271–286.

Bandettini, P.A., Jesmanowicz, J., VanKylen, J., Birn, R.M., Hyde, J.S. 1998. Functional MRI of brain activation induced by scanner acoustic noise. Magnetic Resonance in Medicine 39, 410-416.

Bee, M.A., Klump, G.M. 2004. Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. Journal of Neurophysiology 92. 1088-104.

Benson, D.A., Hienz, R.D. 1978. Single-unit activity in the auditory cortex of monkeys selectively attending to left vs. right ear stimuli. Brain Research 159, 307-320.

Benson, D.A., Hienz, R.D., Goldstein, M.H.J. 1981. Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. Brain Research 219, 249-267.

Boudreau, C.E., Williford, T.H., Maunsell, J.H.R. 2006. Effects of task difficulty and target likelihood in area V4 of macaque monkeys. Journal of Neurophysiology 96, 2377-2387.

Buckner, R.L., Vincent, J.L. 2007. Unrest at rest: Default activity and spontaneous network correlations. Neuroimage 37, 1091-1096.

Cherry, E.C. 1953. Some experiments on the recognition of speech, with one and with two ears. The Journal of the Acoustical Society of America 25, 975-979.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P. 1995. Automated multi-modality registration based on information theory. In: Marchal, G., Bizais, Y., Barillot, C., Di Paolo, R., (Eds.), Proceedings of information processing in medical imaging. Kluwer Academic. pp. 263-274.

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.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., Petersen, S.E. 1991. Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. Journal of Neuroscience 11, 2383-2402.

Cox, R.W., Jesmanowicz, A. 1999. Real-time 3D image registration for functional MRI. Magnetic Resonance in Medicine 42, 1014-1018.

Degerman, A., Rinne, T., Salmi, J., Salonen, O., Alho, K. 2006. Selective attention to sound location or pitch studied with fMRI. Brain Research 1077, 123-134.

Desimone, R., Duncan, J. 1995. Neural mechanisms of selective visual attention. Annual Review of Neuroscience 18, 193-222.

Diamond, D.M., Weinberger, N.M. 1986. Classical conditioning rapidly induces specific changes in frequency receptive fields of single neurons in secondary and ventral ectosylvian auditory cortical fields. Brain Research 372, 357–360.

Diamond, D.M., Weinberger, N.M. 1989. Role of context in the expression of learning-induced plasticity of single-neurons in auditory cortex. Behavioral Neuroscience 103, 471–494.

Downar, J., Crawley, A.P., Mikulis, D.J., Davis, K.D. (2001). A multimodal cortical network for the detection of changes in the sensory environment. Nature Neuroscience 3(3), 277-283.

Duncan, J., Owen, A.M. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. Trends in Neurosciences 23, 475-483.

Duvernoy, H.M. 1999. The human brain: surface, blood supply, and three-dimensional anatomy. 2nd ed. Springer, Wien New York.

Edeline, J.M., Weinberger, N.M. 1993. Receptive field plasticity in the auditory cortex during frequency discrimination training: selective retuning independent of task difficulty. Behavioral Neuroscience 107, 82-103.

Edeline, J.M., Pham, P., Weinberger, N.M. 1993. Rapid development of learning induced receptive field plasticity in the auditory cortex. Behavioral Neuroscience 107, 539-551.

Fishman, Y.I., Reser, D.H., Arezzo, J.C., Steinschneider, M. 2001. Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hearing Research 151,167-187.

Fishman, Y.I., Arezzo, J.C., Steinschneider, M. 2004. Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. Journal of the Acoustical Society of America 116, 1656-1670.

Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., & Goebel, R. 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron, 40(4), 859-869.

Fritz, J.B., Elhilali, M., Shamma, S.A. 2005. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. Journal of Neuroscience 25, 7623-7635.

Fritz, J.B., Elhilali, M., Shamma, S.A. 2007a. Adaptive changes in cortical receptive fields induced by attention to complex sounds. Journal of Neurophysiology 98, 2337-2346.

Fritz, J.B., Elhilali, M., David, S.V., Shamma, S.A. 2007b. Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? Hearing Research 229, 186-203.

Fritz, J.B., Shamma, S.A., Elhilali, M., Klein, D.J. 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nature Neuroscience 6, 1216-1223.

Genovese, C.R., Lazar, N.A., Nichols, T.E. 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15, 870-878.

Griffiths, T.D., Warren, J.D. 2002. The planum temporale as a computational hub. Trends in Neurosciences 25, 348-353.

Gutschalk, A., Micheyl, C., Melcher, J.R., Rupp, A., Scherg, M., Oxenham, A.J. 2005. Neuromagnetic correlates of streaming in human auditory cortex. Journal of Neuroscience 25(22), 5382–5388.

Gusnard, D.A., Raichle, M.E. 2001. Searching for a baseline: Functional imaging and the resting human brain. Nature Reviews Neuroscience 2, 685-694.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W. 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., Bowtell R.W. 2000. Stimulus modulation and task effects in auditory processing measured using fMRI. Human Brain Mapping 10(3),107-119.

Hall, D.A., Plack, C.J. 2009. Pitch processing sites in the human auditory brain. Cerebral Cortex 19(3), 543-553.

Harel, N., Lee, S.-P., Nagaoka, T., Kim, D.-S., Kim, S.-G. 2002. Origin of negative blood oxygenation level-dependent fMRI signals. Journal of Cerebral Blood Flow and Metabolism 22, 908-917.

Hart, H., Palmer, A.R., Hall, D.A. 2002. Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. Hearing Research 171(1), 177-190.

Hart, H.C., Hall, D.A., Palmer, A.R. 2003. The sound-level-dependent growth in the extent of fMRI activation in Heschl’s gyrus is different for low- and high-frequency tones. Hearing Research 179(1-2), 104-112.

Hocherman, S., Benson, D.A., Goldstein, J.M.H., Heffner, H.E., Hienz, R.D. 1976. Evoked unit activity in auditory cortex of monkeys performing a selective attention task. Brain Research 117, 51-68.

Howard, M.A., Volkov, I.O., Abbas, P.J., Damasio, H., Ollendieck, M.C., Granner, M.A. 1996. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Research 724, 260-264.

Hubel, D.H., Henson, C.O., Rupert, A., Galambos, R. 1959. “Attention” units in the auditory cortex. Science 129, 1279-1280.

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. Neurosci. Lett. 266, 125-128.

Johnson, J.A., Zatorre, R.J. 2005. Attention to simultaneous unrelated auditory and visual events: behavioural and neural correlates. Cerebral Cortex 15, 1609-1620.

Johnson, J.A., Zatorre, R.J. 2006. Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. NeuroImage 31, 1673-1681.

Krumbholz, K., Eickhoff, S.B., Fink, G.R. 2007. Feature- and object-based attentional modulation in the human auditory 'where' pathway. Journal of Cognitive Neuroscience 19, 1721-1733.

Logothetis, N.K., & Wandell, B.A. 2004. Interpreting the BOLD signal. Annual Review of Physiology, 66, 735-769.

Luck, S.J., Chelazzi, L., Hillyard, S.A., Desimone, R. 1997. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. Journal of Neurophysiology 77, 24-42.

Micheyl C., Tian B., Carlyon R.P., Rauschecker, J.P. 2005. Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 48, 139-48.

O'Craven, K.M., Downing, P.E., Kanwisher, N. 1999. fMRI evidence for objects as the units of attentional selection. Nature 401, 584-587.

Ohl, F.W., Scheich, H. 1996. Differential frequency conditioning enhances the spectral contrast sensitivity of units in auditory cortex (field A1) of the alert Mongolian gerbil. European Journal of Neuroscience 8, 1001–1017.

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.

Pfingst, B.E., O'Connor, T.A., Miller, J.M. 1977. Response plasticity of neurons in auditory cortex of the rhesus monkey. Experimental Brain Research 29, 393-404.

Price,C.J., Moore, C.J., Friston, K.J. 1997. Subtractions, conjunctions, and interactions in experimental design of activation studies. Human Brain Mapping 5, 264-272.

Schoenfeld, M.A., Hopf, J.M., Martinez, A., Mai, H.M., Sattler, C., Gasde, A., Heinze, H.J., Hillyard, S.A. 2007. Spatio-temporal analysis of feature-based attention. Cerebral Cortex 17, 2468-2477.

Shmuel, A., Augath, M., Oeltermann, A., Logothetis, N.K. 2006. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. Nature Neuroscience 9, 569-577.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., Petersen, S.E. 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. Journal of Cognitive Neuroscience 9, 648-663.

Snyder, J.S., Alain, C., Picton, T.W. 2006. Effects of attention on neuroelectric correlates of auditory stream segregation. Journal of Cognitive Neuroscience 18(1), 1-13.

Somers, D.C., Dale, A.M., Seiffert, A.E., Tootell, R.B.H. 1999. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. Proceedings of the National Academy of Sciences of the United States of America 96, 1663-1668.

Sridharan, D., Levitin, D.J., Menon, V. 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proceedings of the National Academy of Sciences of the United States of America 105(34), 12569–12574.

Studholme, C., Hill, D., Hawkes, D. 1998. A normalized entropy measure of 3-D medical image alighment, Proceedings of Medical Imaging, Vol. 3338, San Diego, CA. pp. 132-143.

Talavage, T.M., Ledden, P.J., Benson, R.R., Rosen, B.R., Melcher, J.R. 2000. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. Hear. Res. 150, 225-244.

Talavage, T.M., Sereno, M.I., Melcher, J.R., Ledden, P.J., Rosen, B.R., Dale, A.M. 2004. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. Journal of Neurophysiology 91, 1282-1296.

Tootell, R.B.H., Hadjikhani, N., Hall, E.K., Marrett, S., Vanduffel, W., Vaughan, J.T., Dale, A.M. 1998. The retinotopy of visual spatial attention. Neuron 21, 1409-22.

Treue, S., Trujillo, J.C.M. 1999. Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399, 575-579.

Wilson, E.C., Melcher, J.R., Micheyl, C., Gutschalk, A., Oxenham, A.J. 2007. Cortical fMRI activation to sequences of tones alternating in frequency: Relationship to perceived rate and streaming. J Neurophysiol 97, 2230–2238.

Table legends

Table 1 Summary of the main statistical contrasts used for testing the frequency-dependent response, and for testing the effects of enhancement and suppression on the frequency-dependent response during the attention-demanding task.

Table 2 The number of suprathreshold ($p<0.01$) voxels showing combinations of the three main effects (frequency sensitivity, attention-specific and general enhancement). Voxel counts are reported separately for the six participants and averaged across the group. Note that data are pooled across left and right hemispheres.

Table 3 The peaks of maximum incidence for two of the putative effects of selective attention (general enhancement and general suppression) for the low- and the high-frequency contrasts. Note that these peaks were observed in incidence maps that were generated from the individual contrast maps, FDR corrected ($p<0.05$). Abbreviations : IFG : inferior frontal gyrus, TPOJ : temporo-parietal-occipital junction, CC : cingulate cortex.

Figure legends

Figure 1 Schematic diagram to represent the two concurrent sound sequences. Panel A) represents the stimulus condition with the low-frequency majority stream. Panel B) represents the stimulus condition with the high-frequency majority stream.

Figure 2 Panel A: The chosen orientation of the functional scan was parallel to the Sylvian fissure. This position is denoted by the black rectangle. Note that the field of view is smaller than the head and so to reduce the resulting Nyquist artefacts saturation bands, marked by the white rectangles, were placed at the anterior and posterior edges of the scan. Panel B: An example of the resultant functional scan showing the nulled MR signal at the temporal poles (top of the image) and the occipital lobe (bottom of the image). Panel C: The region encompassing the superior temporal gyrus and superior temporal sulcus (shown in white) defined the anatomical region of the interest for examining the incidence maps.

Figure 3 Target-detection performance (A) during training (B) in the MR scanner, for the four 'attend' conditions. The height of the bars indicate the average performance for all six participants, the symbols indicate individual performance. Panel (C) shows target-detection performance for the selective attention and the divided attention task for participants #4, #5 and #6 tested separately. L: Low-frequency majority stimulus. H: high-frequency majority stimulus.

Figure 4 Distribution of frequency sensitivity, attention-related enhancement, general enhancement and general suppression shown as an incidence map of activity for the six listeners ($p<0.01$). The three horizontal slices are chosen to illustrate activity throughout the auditory cortex and the spread of activation can be viewed relative to the position of HG (outlined in white), and PT (outlined in yellow). The images in the left-hand column show the results for the low-frequency contrasts, and the images in the right-hand column show the results for the high-frequency contrasts.

Figure 5 Venn diagram displaying the number of suprathreshold ($p<0.01$) voxels showing the three main effects (frequency-dependence, general and attention-specific enhancement), separately for low- and high-frequency contrasts. Voxel counts show the average across the six participants using the same values that are reported in Table 2. The area of the circles represents the proportional relationship between the extent of the different effects.

Figure 6 A schematic summary of the different regions of functional activity discussed in the text. Some regions are located on the cortical surface and these are shown on the lateral view. Other regions are located in the inner surface and these are shown on the medial view. The insular cortex is tucked away deeply in the Sylvian fissure.

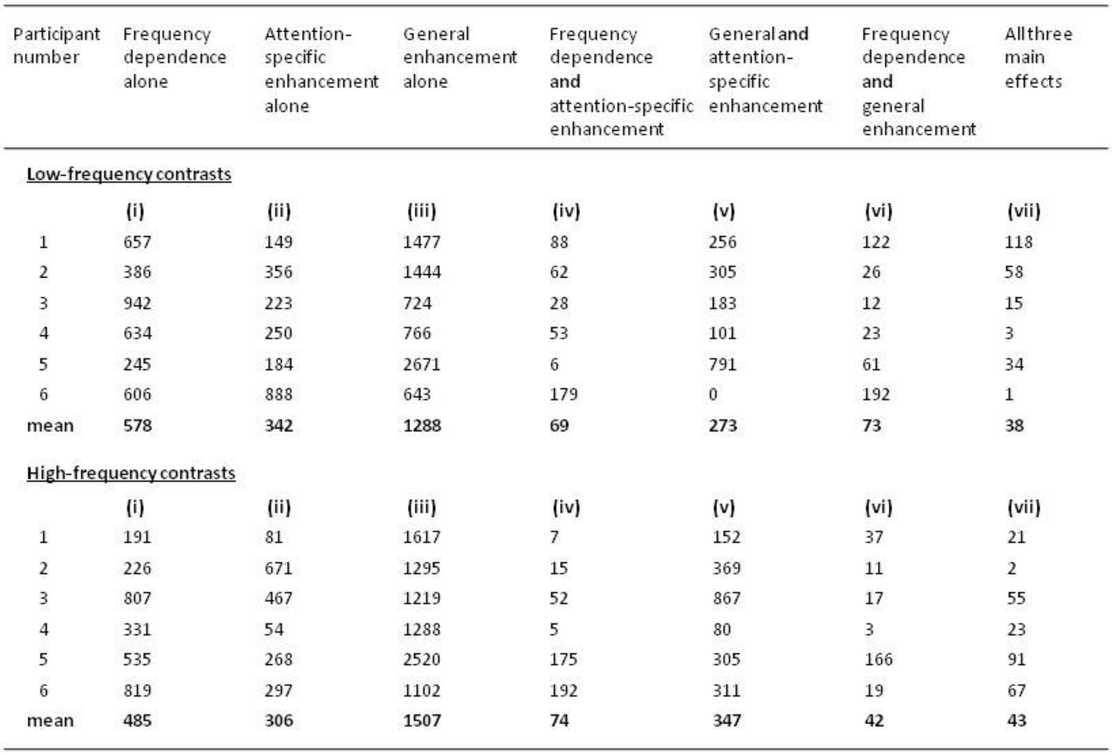
Figure 7 Distribution of general suppression in TPOJ, precuneus and posterior cingulate cortex for the low-frequency contrast. This is shown as an incidence map displayed in sagittal and axial views for the six listeners, based on the individual FDR corrected ($p<0.05$) maps. The graphs for each region of interest show the average effect size for the voxels of peak incidence reported in Table 3. The error bars denote the 95%

confidence intervals. L: Low-frequency majority stimulus. H: high-frequency majority stimulus.

Table 1 Summary of the main statistical contrasts used for testing the frequency-dependent response, and for testing the effects of enhancement and suppression on the frequency-dependent response during the attention-demanding task.

Statistical contrast		Functional response
Just listen low-frequency majority stimulus > Just listen high-frequency majority stimulus	A>B	Low-frequency dependent response
Just listen high-frequency majority stimulus > Just listen low-frequency majority stimulus	B>A	High-frequency dependent response
Attend low-frequency targets in the low-frequency majority stimulus > Attend high-frequency targets in the low-frequency majority stimulus	C>E	Attention-specific enhancement by selectively attending to low-frequency stream
Attend high-frequency targets in the high-frequency majority stimulus > Attend low-frequency targets in the high-frequency majority stimulus	F>D	Attention-specific enhancement by selectively attending to high-frequency stream
Attend low-frequency targets in the low-frequency majority stimulus > Just listen low-frequency majority stimulus	C>A	General enhancement by selectively attending to low-frequency stream
Attend high-frequency targets in the low-frequency majority stimulus > Just listen high-frequency majority stimulus	F>B	General enhancement by selectively attending to high-frequency stream
Attend high-frequency targets in the low-frequency majority stimulus < Just listen low-frequency majority stimulus	E<A	General suppression by attending away from low-frequency stream
Attend low-frequency targets in the high-frequency majority stimulus < Just listen high-frequency majority stimulus	D<B	General suppression by attending away from high-frequency stream

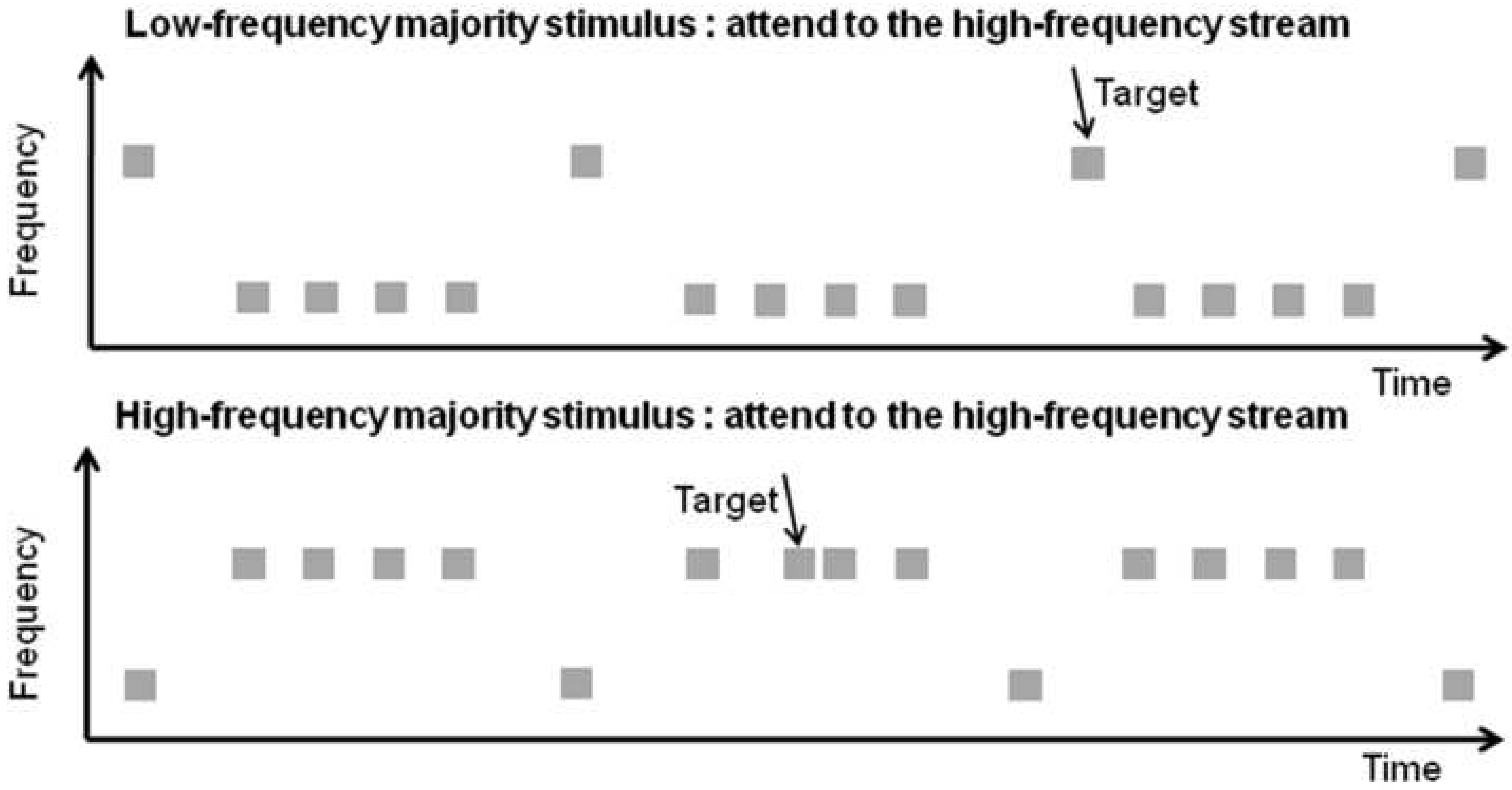
Table 2 The number of suprathreshold (p<0.01) voxels showing combinations of the three main effects (frequency sensitivity, attention-specific and general enhancement). Voxel counts are reported separately for the six participants and averaged across the group. Note that data are pooled across left and right hemispheres.

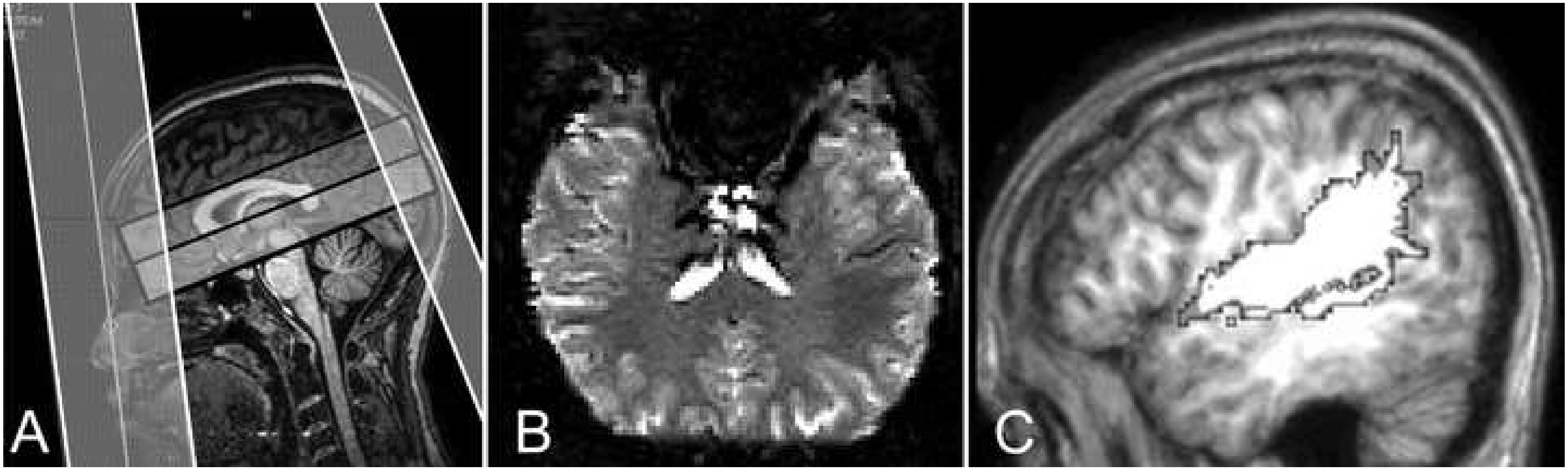


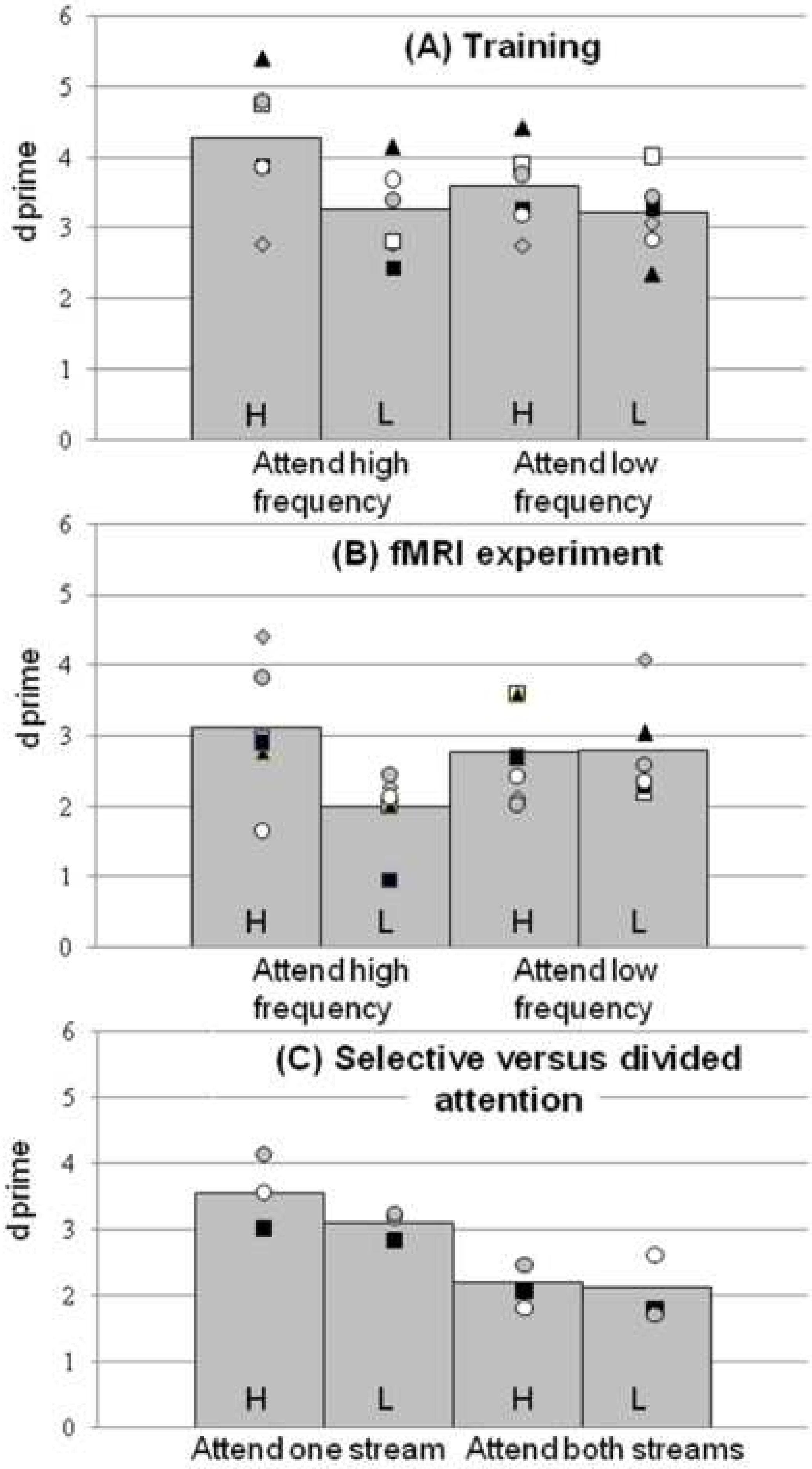
Participant number	Frequency dependence alone	Attention-specific enhancement alone	General enhancement alone	Frequency dependence and attention-specific enhancement	General and attention-specific enhancement	Frequency dependence and general enhancement	All three main effects
Low-frequency contrasts							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
1	657	149	1477	88	256	122	118
2	386	356	1444	62	305	26	58
3	942	223	724	28	183	12	15
4	634	250	766	53	101	23	3
5	245	184	2671	6	791	61	34
6	606	888	643	179	0	192	1
mean	**578**	**342**	**1288**	**69**	**273**	**73**	**38**
High-frequency contrasts							
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
1	191	81	1617	7	152	37	21
2	226	671	1295	15	369	11	2
3	807	467	1219	52	867	17	55
4	331	54	1288	5	80	3	23
5	535	268	2520	175	305	166	91
6	819	297	1102	192	311	19	67
mean	**485**	**306**	**1507**	**74**	**347**	**42**	**43**

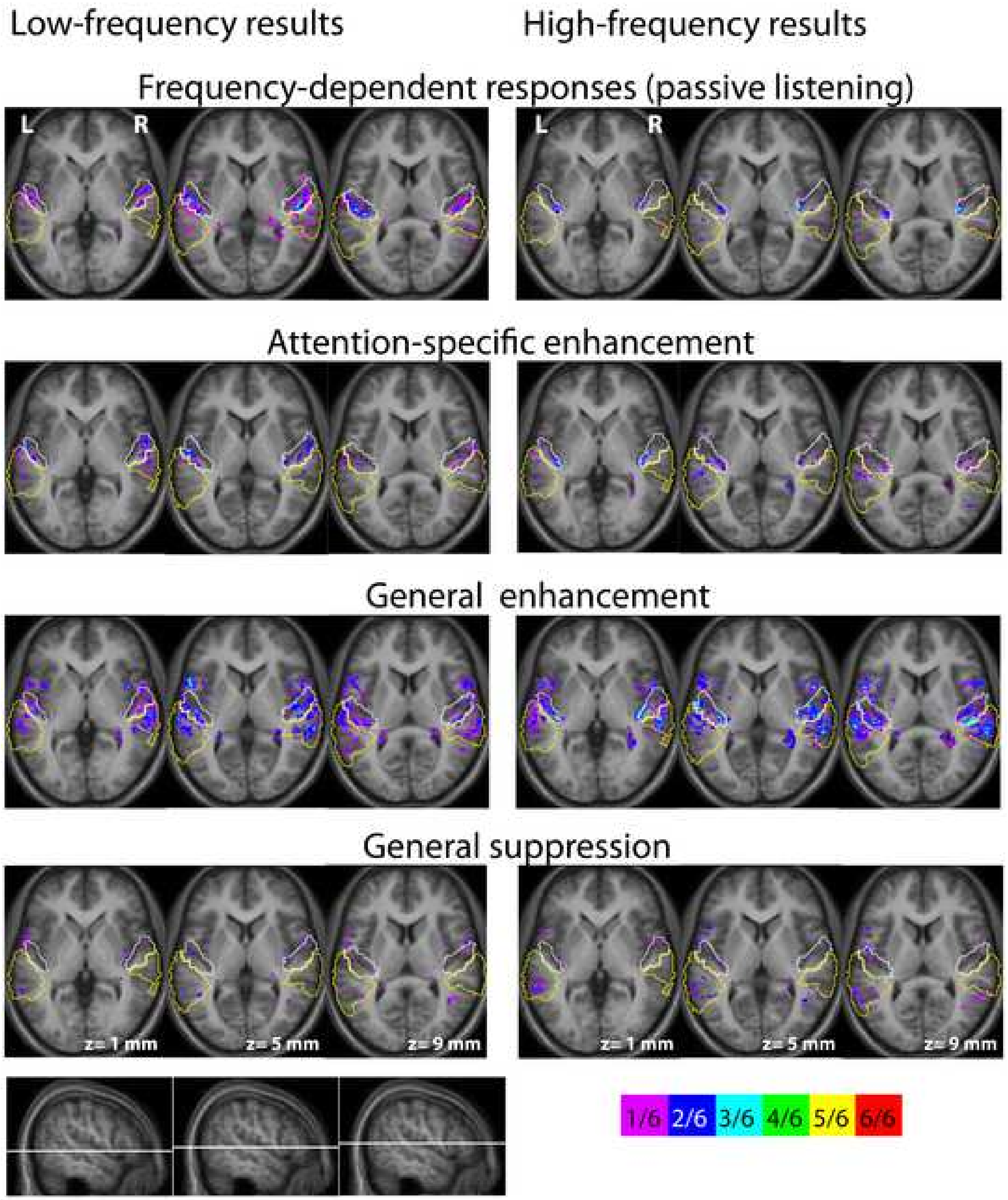
Table 3 The peaks of maximum incidence for two of the putative effects of selective attention (general enhancement and general suppression) for the low- and the high-frequency contrasts. Note that these peaks were observed in incidence maps that were generated from the individual contrast maps, FDR corrected (p<0.05). Abbreviations : IFG : inferior frontal gyrus, TPOJ : temporo-parietal-occipital junction, CC : cingulate cortex.

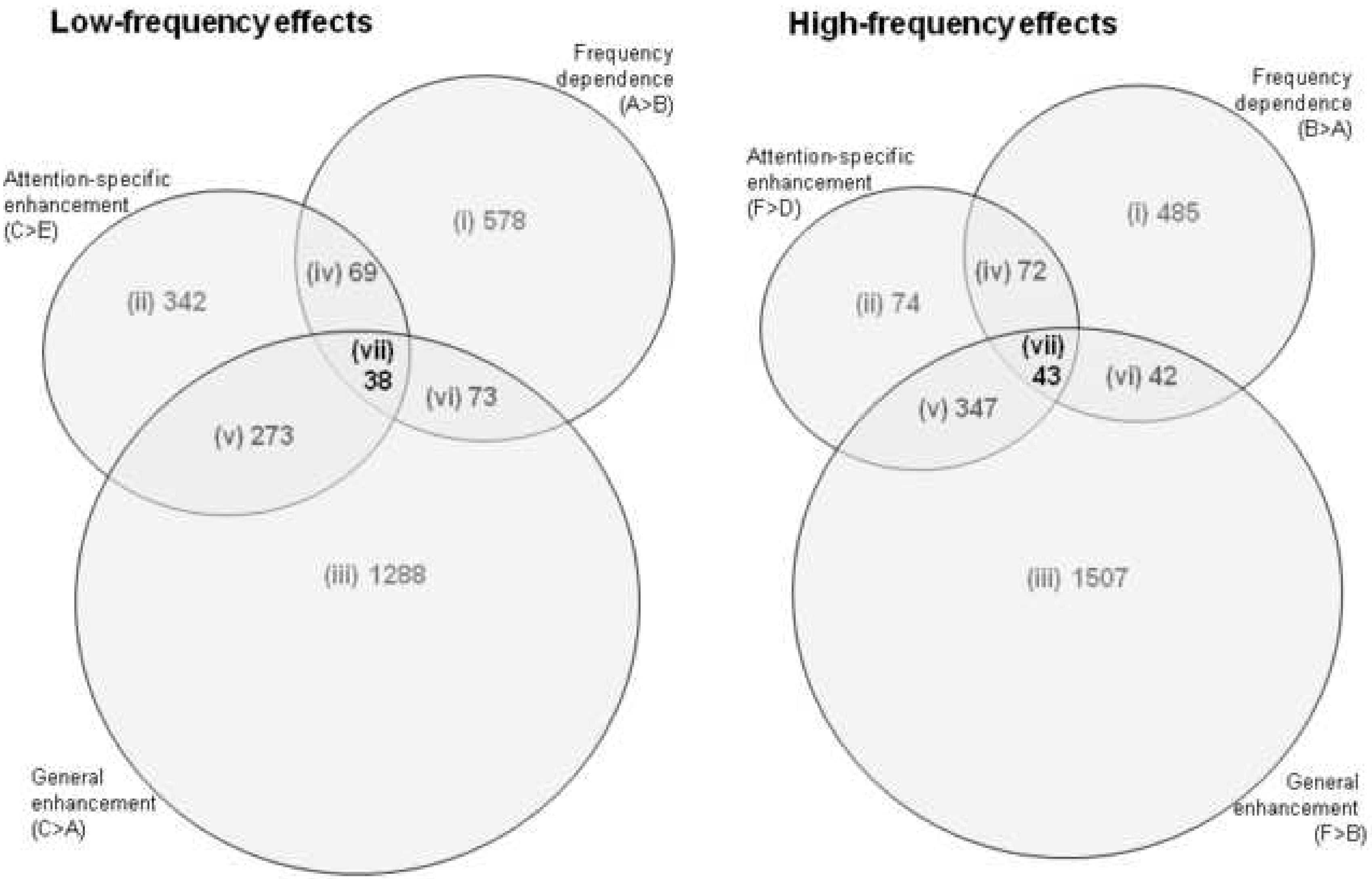
	Low-frequency contrast				High-frequency contrast			
	Maximum incidence	x	y	z	Maximum incidence	x	y	z
General enhancement								
Left IFG	4	-56	14	5	4	-52	6	12
Left insula	4	-32	22	2	4	-32	26	12
Right insula	4	33	22	0	5	33	24	0
General suppression								
Left TPOJ	4	-42	-60	32	4	-44	-57	28
Right TPOJ	5	48	-57	25	4	50	-57	22
Left precuneus	4	-6	-56	40				
Right precuneus	4	2	-50	48	3	8	-48	50
Posterior CC	4	0	-50	32	4	0	-48	30

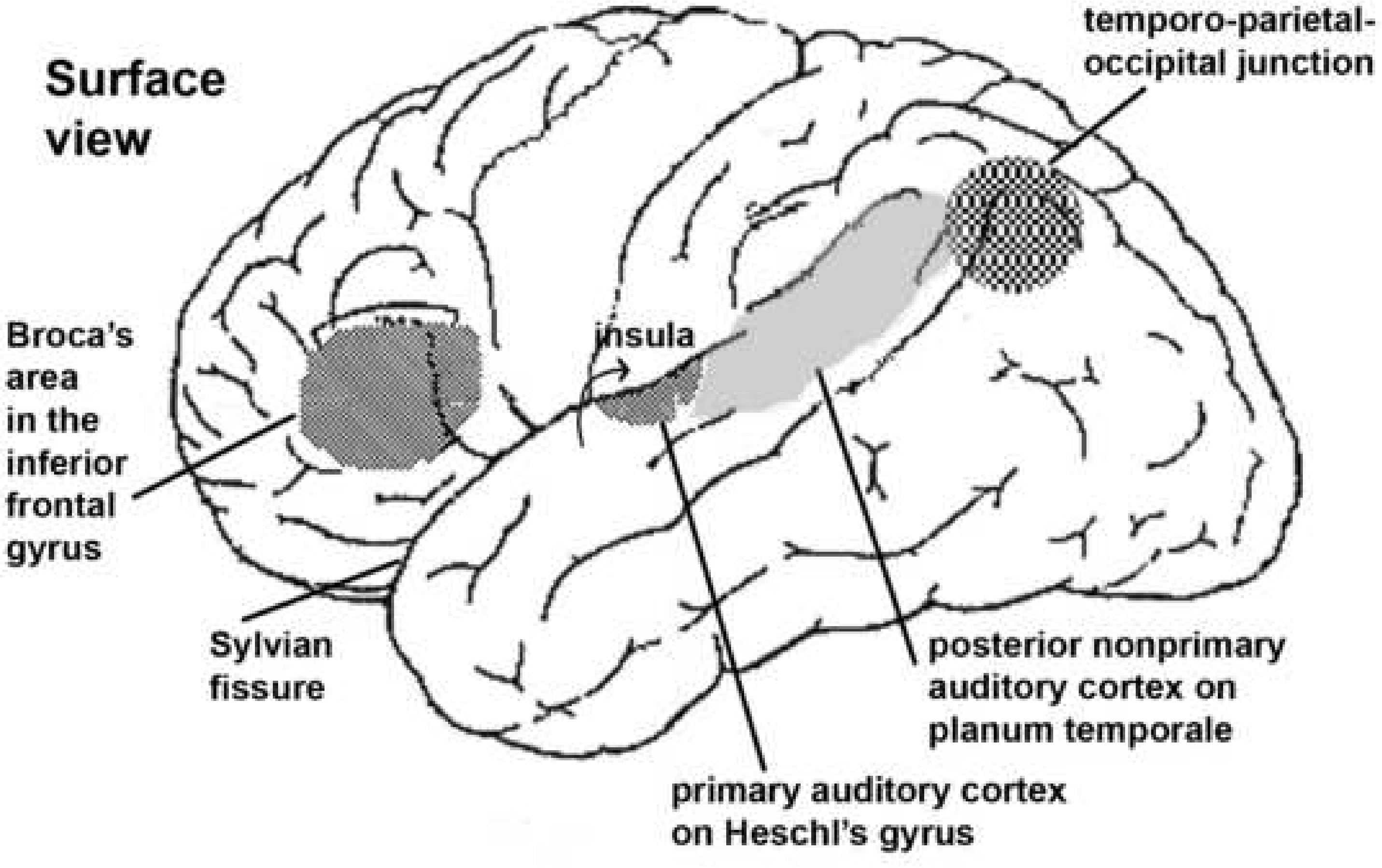

Low-frequency majority stimulus : attend to the high-frequency stream
Target
Frequency
Time
High-frequency majority stimulus : attend to the high-frequency stream
Target
Frequency
Time


A
B
C

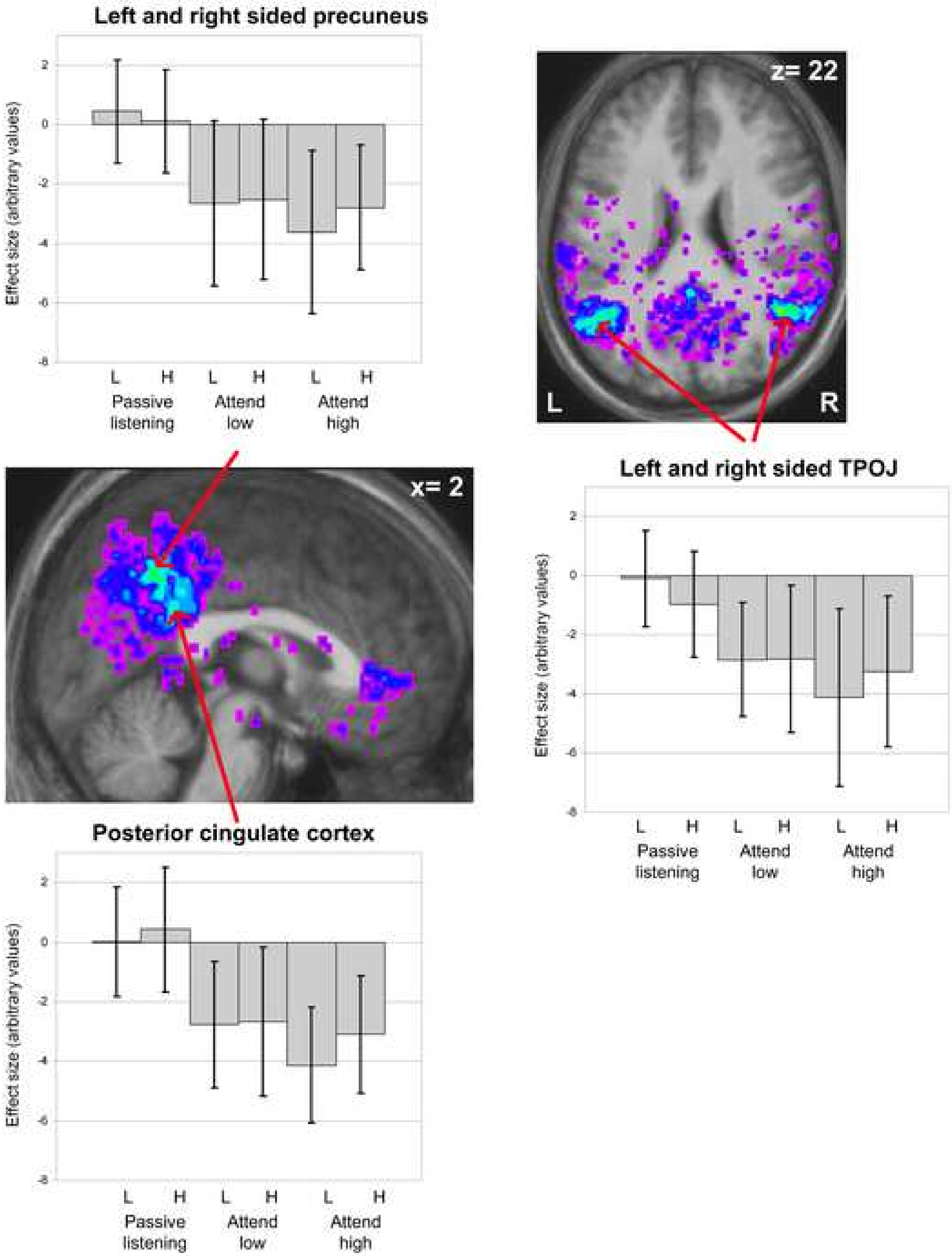

(A) Training
d prime
6
5
4
3
2
1
0
H
L
H
L
Attend high frequency
Attend low frequency
(B) fMRI experiment
d prime
6
5
4
3
2
1
0
H
L
H
L
Attend high frequency
Attend low frequency
(C) Selective versus divided attention
d prime
6
5
4
3
2
1
0
H
L
H
L
Attend one stream
Attend both streams


Low-frequency results
High-frequency results
Frequency-dependent responses (passive listening)
L
R
L
R
Attention-specific enhancement
General enhancement
General suppression
z= 1 mm
z= 5 mm
z= 9 mm
z= 1 mm
z= 5 mm
z= 9 mm
1/6
2/6
3/6
4/6
5/6
6/6


Low-frequency effects
Frequency dependence (A>B)
Attention-specific enhancement (C>E)
(i) 578
(ii) 342
(iv) 69
(vii) 38
(vi) 73
(v) 273
(iii) 1288
General enhancement (C>A)
High-frequency effects
Frequency dependence (B>A)
Attention-specific enhancement (F>D)
(i) 485
(iv) 72
(ii) 74
(vii) 43
(vi) 42
(v) 347
(iii) 1507
General enhancement (F>B)


Surface view
temporo-parietal-occipital junction
Broca's area in the inferior frontal gyrus
insula
Sylvian fissure
posterior nonprimary auditory cortex on planum temporale
primary auditory cortex on Heschl's gyrus


Medial view
precuneus
posterior cingulate cortex


Left and right sided precuneus
Effect size (arbitrary values)
2
0
-2
-4
-6
-8
L
H
L
H
L
H
Passive listening
Attend low
Attend high
z= 22
L
R
x= 2
Left and right sided TPOJ
Effect size (arbitrary values)
2
0
-2
-4
-6
-8
L
H
L
H
L
H
Passive listening
Attend low
Attend high
Posterior cingulate cortex
Effect size (arbitrary values)
2
0
-2
-4
-6
-8
L
H
L
H
L
H
Passive listening
Attend low
Attend high