Title: Acoustic, psychophysical and neuroimaging measurements of the effectiveness of active cancellation during auditory functional magnetic resonance imaging.

Running title: Noise reduction during auditory neuroimaging

* Deborah A. Hall (d.hall@ihr.mrc.ac.uk)
MRC Institute of Hearing Research, University Park, Nottingham, UK. NG7 2RD
Tel: +44 (0)115 9223431
Fax: +44 (0)115 9518503

John Chambers (john.chambers@intelligent-energy.com)
MRC Institute of Hearing Research, University Park, Nottingham, UK. NG7 2RD
Tel: +44 (0)1509 225863
Fax: +44 (0)1509 223911

Michael A Akeroyd (maa@ihr.gla.ac.uk)
MRC Institute of Hearing Research (Scottish Section), Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow, UK. G31 2ER
Tel: +44 (0)141 2114695
Fax: +44 (0)141) 5528411

John R Foster (john@ihr.mrc.ac.uk)
MRC Institute of Hearing Research, University Park, Nottingham, UK. NG7 2RD
Tel: +44 (0)115 9223431
Fax: +44 (0)115 9518503

Ron Coxon (Ron.Coxon@nottingham.ac.uk)
Sir Peter Mansfield Magnetic Resonance Centre, School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, UK. NG7 2RD
Tel: +44 (0)115 9514747
Fax: +44 (0)115 9515166

Alan R. Palmer (alan@ihr.mrc.ac.uk)
MRC Institute of Hearing Research, University Park, Nottingham, UK. NG7 2RD
Tel: +44 (0)115 9223431
Fax: +44 (0)115 9518503

* corresponding author

ABSTRACT

Functional magnetic resonance imaging (fMRI) is one of the principal neuroimaging techniques for studying human audition, but, paradoxically, it generates an intense background sound which hinders listening performance and confounds measures of the auditory response. This paper reports the perceptual effects of an active noise control (ANC) system that operates in the electromagnetically hostile and physically compact neuroimaging environment to provide significant noise reduction, without interfering with image quality. Cancellation was first evaluated at 600 Hz, corresponding to the dominant peak in the power spectrum of the background sound and hence at which cancellation is maximally effective. Microphone measurements at the ear demonstrated 35 dB of acoustic attenuation (from 93 to 58 dB SPL), while masked detection thresholds improved by 20 dB (from 74 to 54 dB SPL). Considerable perceptual benefits were also obtained across other frequencies, including those corresponding to dips in the spectrum of the background sound. Cancellation also improved the statistical detection of sound-related cortical activation, especially for sounds presented at low intensities. These results confirm that ANC offers substantial benefits for fMRI research.

I. INTRODUCTION

In recent years functional magnetic resonance imaging (fMRI) has become a popular technique for non-invasive measurements of central auditory function in humans (e.g. Hall *et al.*, 2003; Hall 2006; Johnsrude *et al.*, 2002). Its use, however, is complicated by the problem that the particular MR pulse sequence —echo-planar imaging (EPI)— used to acquire the images of functional brain activity also generates an intense acoustic noise [1]. This noise results from the switching of currents in the gradient coil of the MR scanner. Because this occurs in a large static magnetic field, it effectively turns the MR scanner into a giant loudspeaker. The noise level typically exceeds 100 dB SPL at the position of a listener (Foster *et al.*, 2000; More *et al.*, 2006; Ravicz *et al.*, 2000). To protect listeners against the damaging noise exposure during scanning, passive noise reduction is routinely achieved by ensuring that listeners wear earplugs and/or circumaural ear defenders (Ravicz and Melcher, 2001). When these methods are used together, a considerable amount of attenuation can be achieved. For example, the ear defenders used in the present set of experiments provide 25 dB of attenuation at 600 Hz rising to 40 dB at 4 kHz. However, passive attenuation alone is still insufficient for acceptably quiet conditions, and even with it, the distracting, loud, repetitive noise is known to elevate feelings of anxiety, emotional distress (Quirk *et al.*, 1989) and claustrophobia (Granet and Gelber, 1990). It has also been shown to interfere with perceptual judgements (e.g. Boyle *et al.*, 2006). Of particular concern for fMRI studies of sound processing is the masking of the background noise on stimuli presented to the listener through headphones and its interference with the measures of stimulus-related brain

activation (Elliott *et al.*, 1999; Shah *et al.*, 1999). Moreover, EPI noise itself produces measurable brain responses not only within the primary auditory cortex (Bandettini et al., 1998), but also within surrounding non-primary auditory regions (Hall et al., 2000; Ulmer et al., 1998). It is therefore preferable to reduce the noise level by other means.

Three different approaches to noise reduction have been reported. One is to reduce the noise at source, but since the source of the sound — the gradient coil switching — is an essential part of the scanning process, the technical challenges in reducing the induced mechanical vibrations of the scanner are considerable. A second approach has been to develop scanning protocols that minimize the instantaneous level or the total dosage of the background noise, but these either require specialized pulse programming (Hennel *et al.*, 1999; Schwarzbauer *et al.*, 2006) or dramatically increase the scanning time required to obtain the same number of brain images (Hall *et al.*, 1999); for a review see Amaro *et al.* (2002). A third approach is to acoustically reduce the level of the sound using the technique of active noise control (ANC). This technique is examined here.

*** Insert Figure 1 about here ***

The spectral characteristics of the EPI noise are mainly a function of the "readout" gradient pulses of the chosen MR pulse sequence and comprise a series of harmonically related peaks together with a background of broadband noise. The intensity of the peaks is primarily attributable to the features of the gradient-coil assembly and so can differ across manufacturers and models (More *et al.*, 2006). The peak level of the EPI noise produced by the MR scanner that was used in the current series of experiments is 99 dB SPL measured at the center of the bore. The dark line in Figure 1 illustrates the spectrum of this noise. It has a peak at 600 Hz with

two other prominent peaks, 5-10 dB smaller, at 150 and 300 Hz. The harmonic structure of the noise, and its spectral invariance, make it an ideal candidate for active cancellation. Moreover, the noise repeats at an exact period determined by the EPI pulse sequence. In the example shown in Figure 1 (see inset), there is a rapid succession of 18 'pings', corresponding to a series of 18 consecutive slices acquired through the brain. The repeating period is 64 ms and the whole noise period lasts approximately 1.15 seconds. This temporal predictability is also critical to the success of ANC.

One ANC approach that has been taken is to acoustically control the gradient coils themselves, a method which has been shown to greatly reduce the noise generated at the source within the coil structure (Chapman *et al.* , 2003). However, so far this method amounts only to a technical demonstration and thus it still requires incorporation into commercial MR systems for the benefits to be realized in practice. A second approach that we have taken has been to design an active acoustically controlled headset. This method uses a cancellation signal to add destructively with the noise at the listener's ear. It has been shown to greatly reduce the noise at the eardrum of the listener and poses no restrictions on the MR scanning hardware (Chambers *et al.*, 2001; 2007). Previously, we have reported the technical specification of a two-channel sound system for the active control of EPI noise that worked in conjunction with passive attenuation (using circumaural ear defenders) (Chambers *et al.*, 2001) and subsequent modifications to optimise the performance of this system (Chambers *et al.*, 2007). This prototype system was originally evaluated in the laboratory using a plastic-and-wood replica of a scanner in which recordings of EPI noise were presented over loudspeakers. For both recorded examples of noise and for simulated scanner sounds with dominant frequencies from 500 Hz to 3 kHz, the ANC

system reduced the levels of the peak components by about 40 dB, when measured with a microphone placed underneath the circumaural ear defenders. The perceptual benefits of noise reduction were ascertained by making loudness matches between cancelled and non-cancelled targets that were created from recordings of EPI noise that had been filtered to remove energy at the frequencies remote from the fundamental (here the dominant components were 600 and 1900 Hz, respectively). In contrast to the acoustic cancellation, perceptual cancellation reached, at most, only 13 dB. We proposed that the difference between acoustic and perceptual measures of noise reduction was likely to be due to the rise of the bone conduction route as the primary path of sound to the cochlea when attenuation of the airborne sound had reached its limit. Accordingly, Chambers *et al.* (2001) predicted that ANC should be most effective when the most intense component of the EPI noise occurs at low frequencies (< 1000 Hz), where the difference between bone and air conduction is greatest (Berger, 1983)[2].

The main goal of the three experiments reported here was to examine the effectiveness of this ANC system under true fMRI conditions. We used a Philips 3 T MR scanner for acquiring EPI scans during all of our tests. The noise generated by this MR scanner contains three prominent peaks below 1 kHz (Figure 1) and is representative of the EPI noise generated by other commercial MR scanners that are currently available. This environment is well suited for testing the ANC system because it operates within the range of peak frequencies at which perceptual benefits might be greatest. However, body vibrations are much greater in a real MR scanner than in a simulation, through the coupling between the bed and the gradient coil system. The unknown characteristics of the bone-transmitted energy mean that it is uncertain how much perceptual benefit might be provided by ANC. Both Experiments 1 and 3 investigated the

perceptual benefits of the reduction in the EPI noise masking by measuring masked detection thresholds of targets presented at different frequencies, with and without the ANC system operating. In contrast, Experiment 2 evaluated the effects of ANC on the detectability of stimulus-related activation measured by fMRI using sounds whose frequency was close to the dominant peak at 600 Hz. In all three experiments the EPI noise itself served as the masker.

An important component of the ANC system is the adaptive filter which optimizes the cancellation signal. The coefficients of this filter are determined during an initial 'training' period, but they can then be either fixed for the remainder of the test or allowed to continually adjust. Previously, we used a fixed-coefficient ANC system (Chambers *et al.*, 2001; 2007), but if either the spectral content of the scanner noise or the position of the listener's head changes sufficiently over time, then the effectiveness of cancellation will be diminished. Head movement in particular affects the phase of the incident sound wave, and so is of particular concern. Given that it takes several minutes to arrive at the detection threshold by adaptive procedures, fixing the cancellation signal may pose a real concern for the practical implementation of the system and, in principle, compromise the operation of the ANC system. Accordingly, Experiments 1 and 2 were conducted using an ANC system in which the filter coefficients were continually updated. For comparison, Experiment 3 was conducted using an ANC system in which the filter coefficients were fixed.

II. EXPERIMENT 1: EFFECTS OF CANCELLATION AT THE PROMINENT FREQUENCY OF THE EPI NOISE

We measured the perceptual benefit of the ANC system psychophysically by measuring detection thresholds for a target consisting of a narrowband noise when masked by the EPI noise. Henceforth when used in this context, we refer to the EPI noise as the masker. In the "cancelled" condition, the ANC system attenuated the masker. In the "non-cancelled" condition, the ANC system was switched off, and so the EPI noise was unchanged. The mean difference in thresholds between the two conditions provides a measure of the perceptual reduction in noise masking achieved by the system. The level of the target at threshold was also compared with the level of the residual masker noise, measured using a microphone positioned under the ear defender close to the entrance to the ear canal.

A. Methods

1. Description of the ANC system

The present ANC system forms part of a fully integrated user-friendly MR-compatible sound system that enables the presentation of high-fidelity sound stimuli to participants during auditory fMRI research. It uses Sennheiser HE60 electrostatic headphones incorporated into Bilsom 2452 circumaural ear defenders, together with an external high-voltage amplifier (Sennheiser HEV70; see www.ihr.mrc.ac.uk/research/technical.php).

Only a brief summary of the ANC system is given here because more detailed reports of the technical specification can be found in Chambers *et al.* (2001; 2007). The ANC system operates independently in the left and right channels. The system is similar in concept to a standard feed-forward noise control system (Nelson and Elliott, 1992), which has two

microphones, termed the ‘reference’ and ‘error’, and which perform two different functions. The first function (performed by the reference microphone) is to receive the sound that is propagating towards the point of cancellation. This sound is termed the reference signal. The second function (performed by the error microphone) is to monitor the residual sound after cancellation. Ideally, in an MR scanner, the reference microphone would be close to the noise source (i.e. the gradient coils) and the error microphone would be near to the listener’s eardrum. However, the two microphones must be sufficiently separated in space to enable a long enough propagation delay between them for generating the cancellation signal. In the MR scanner, this cannot be achieved because the gradient coils are too close to the listener’s head. Accordingly, our system has only one microphone that serves both functions, but at different times. This microphone uses optical technology so that the radio-frequency energy that is used in the scanning process does not affect its operation, and its presence does not produce any deterioration of the brain images acquired. It is positioned under the ear defender, close to the ear canal. The electrostatic transducer of the headset also has a dual role. It delivers both the experimental sound stimuli, as well as the cancellation signal.

The standard function of the reference microphone is replaced in our system by a pre-recording of the EPI noise that serves as the reference signal. This recording is taken from the microphone during one of the initial bursts of EPI noise immediately before cancellation is started. It cannot be used directly as the cancellation signal because it is affected by the transfer function of the ANC system, including the electrostatic transducer and the error microphone in the headset. Therefore the transfer function is compensated for by a digital filter (termed the ‘plant model’), which is calculated before the start of cancellation by presenting an MLS

sequence through the system and receiving it on the microphone. This is done during the initialization phase of the MR scanner, when the coolant pump (which recycles liquid gases surrounding the superconducting magnet) automatically switches off and before the gradient switching begins[3].

During ANC operation, the cancellation signal is presented through the transducer, so that it adds destructively with the EPI noise at the listener's ear. The level of residual noise at the microphone (now operating as the error microphone) is continually monitored and used to fine tune the cancellation signal. The timing of the cancellation signal is determined by an electrical trigger pulse, produced by the MR scanner 25 ms in advance of the EPI noise. The cancellation signal itself is derived from the reference signal via an adaptive digital filter, whose coefficients are set so that cancellation is as effective as possible (i.e., the resultant signal at the error microphone is as small as possible). The initial setting is found during four 'training' scans and uses a fast rate of change of adaptation. These coefficients are then subsequently adjusted at a slower rate of adaptation, to track slow changes in the characteristics of the EPI noise. This two-step process improves the immunity of the adaptive filter from the stimuli that are being presented during the fMRI experiment and hence helps to maximize the acoustic noise reduction.

In practice, the ANC system is easy to operate, and it is integrated into the same PC software used for presenting the sound stimuli during the fMRI experiment. The three stages involved in setting-up and using the system (determining the plant model, acquiring the reference signal, and optimising the coefficients of the adaptive filter) are all automatically controlled within this software and are specified by only a few parameters.

2. Stimuli

The targets were diotic narrowband noises, centered at 600 Hz, with a bandwidth of 50 Hz, a duration of 500-ms (including 50-ms raised-cosine ramps applied to the onset and offset) and sampled at a rate of 44,100 Hz. Stimuli were first created in the spectral domain (using a fast Fourier transform with 2-Hz resolution) with rectangular edges to the passband, before windowing in the time domain using Matlab. A set of 100 examples of the targets were created in advance, and the total RMS power of each was equated. On each trial, a sound file was chosen randomly from the set as the target. The presentation level of the target was varied according to the adaptive procedure by applying a digital attenuation before digital-to-analogue conversion; 0-dB attenuation corresponded to a level of 94 dB SPL. This, and subsequent, stimulus presentation levels were calibrated by mounting the headset on a KEMAR manikin equipped with a free-field response microphone (Brüel and Kjær, Type 4134) and Zwislocki Coupler (Brüel and Kjær, Type DB-100), and connected to a measuring amplifier (Brüel and Kjær, Type 2636).

The maskers were the noises generated by the MR scanner (a Philips 3 Tesla Intera equipped with an 8-channel receiver head coil) as it acquired sets of EPI brain images. Each noise burst comprised a rapid succession of 18 ‘pings’ occurring over 1.15 seconds (see inset, Figure 1). Each ‘ping’ corresponded to a brain slice and the set of slices was angled in an oblique horizontal plane sufficient to cover the superior temporal gyrus (auditory cortex) in both hemispheres. Note that in Experiments 1 and 3, these brain images were not analysed because only the psychophysical data were relevant to the question. However, it should be emphasised that the scans were proper functional images.

Conventional fMRI has equal intervals between each set of image acquisitions, and so the noises bursts are equally spaced However, this sequence would have caused difficulties for creating the timing of the standard psychophysical method of two closely spaced intervals followed by an inter-trial gap. Accordingly, we used a bespoke EPI pulse sequence to generate a noise sequence that would permit us to conduct a two-interval psychophysical paradigm. In this paradigm, a "trial" comprised a pair of masker noise bursts separated by a gap of 0.85 s. An inter-trial interval of 2.85 s provided a suitable response window. The response window between each pair of masker noise bursts was created by applying a modified gradient waveform comprising only slice-selective excitation when the radio-frequency excitation pulse was applied (see Schwarzbauer *et al.*, 2006). Because this event did not generate an intense noise nor create a brain image, it is termed a "dummy" scan. Note that this bespoke sequence was only required for the psychophysical evaluation and would not be necessary for the more routine implementation of active noise control in auditory fMRI.

3. Procedures

The detection threshold for the target was measured using two-interval, two-alternative forced-choice (2I-2AFC) task allied to a two-down, one-up adaptive procedure that converged upon the 70.7% correct point on the psychometric function (Levitt, 1971). The target was presented randomly within one of the noise maskers and its onset always occurred 0.40 s after the onset of the masker. The listener was required to judge whether the target appeared in the first or second interval, responding appropriately by pressing one of two buttons with the index finger of either their left or right hand. At its initial level, the target was clearly audible (79 dB SPL in the

cancelled condition and 89 dB SPL in the non-cancelled condition). This level was used for the first two trials of each run, classed as practice trials. Thereafter, the level changed according to performance in step sizes of 5 dB for the first two reversals, and then 2 dB for a further four reversals. The mean of the levels at the last four reversal points was taken as the detection threshold for that run. During each adaptive run, the residual sound at the error microphone inside the earcups was recorded onto a DAT recorder (Sony DTC-690) for off-line evaluation of the acoustic cancellation over time. For each listener, ten adaptive runs were completed for both the non-cancelled and the cancelled conditions. Each run lasted about 3 ½ minutes. To avoid fatigue, these measurements were spread across two 1½-hour testing sessions. In each session, five adaptive runs were completed for each condition, the order of conditions being counterbalanced across sessions and listeners in an ABBA manner. Listeners were allowed a short break after each adaptive run.

All listeners had participated in an earlier practice-and-familiarization session using our plastic-and-wood replica of an MR scanner located in a double-walled sound-attenuating room (see Chambers *et al.*, 2001, Appendix B). Listeners lay with their head inside a 'receiver coil' of realistic dimensions surrounded, at a distance of 15-20 cm, by three loudspeakers (Eminence Alpha/6A) through which digital recordings of the EPI noise were presented. Although the level of the simulated EPI noise did not quite reach that measured in the actual scanner bore, the procedures were sufficiently similar for learning the detection task and the use of the response box. Feedback on performance was given only at the end of each run. Listeners completed four non-cancelled adaptive runs in which their thresholds differed by no more than 6 dB.

3. Listeners

Four right-handed listeners participated in Experiment 1. Participants were two males aged 42 and 39 years, and two females aged 18 and 36 years (henceforth referred to as subjects 1-4 respectively). Listeners 2 and 4 were authors of this article. All listeners had normal hearing as assessed using standard pure-tone audiometry (< 20 dB HL over octave frequencies between 250 and 8000 Hz).

B. Analysis

The residual sound that was recorded from the error microphone was analysed to determine the sound level at the error microphone over the entire duration of the cancelled runs, including the training period. Since Experiment 1 is concerned only with the amount of cancellation at the frequency where the EPI noise is most intense, the digitized recordings were first filtered with a 100-Hz-wide filter centered on 600 Hz (Adobe Audition FFT filter) to remove sound energy that was remote from the frequency of interest. We chose a bandwidth of 100 Hz as a close approximation to the equivalent rectangular bandwidth (ERB) of the psychophysical auditory filter at 600 Hz (Glasberg and Moore, 1990). These filtered recordings were then processed using a customized program to calculate the RMS value for each burst of EPI noise. The program detected the portions of the recording in which the residual sound exceeded a certain threshold and it then calculated an RMS value over a central 1-s time window that contained either the masker alone or a combination of the masker and the 600-Hz target. The recorded levels were then calibrated against the known level of the target presented on each trial so that we obtained

an accurate RMS level estimate over time at each error microphone. Note that, due to the 2I-2AFC procedure, for each pair of noise bursts, one value represented the combined sound level for the target and the masker, while the other represented only the masker. The noise levels for dummy scans are not reported. These RMS levels of the masker noise in the cancelled condition are highly informative because the difference between the levels of the masker noise at the start of the run (corresponding to the first "recording" scan) and at the end portion of the adaptive run (corresponding to the psychophysical threshold reversal points) defines the amount of acoustic cancellation achieved by the ANC system. The absolute values also provide information about differences across listeners and across ears (within a listener). We also compared the values of the masker to the presentation level of the target on the corresponding trial, for the last four threshold reversal points, to investigate whether the level of the masker was the limiting factor for perception. Masker levels were measured using only those intervals that did *not* contain the 600-Hz target and are reported separately for the better and worse ear since the amount of acoustic noise reduction was not always equal at the two ears. It is reasonable to assume that listeners performed the detection task using their better ear, irrespective of whether this was the left or the right.

C. Results

1. Acoustic noise reduction during the cancelled condition

Table I shows the non-cancelled and cancelled masker levels for each listener and on average these were 93 and 58 dB SPL respectively. Thus, the ANC system provided 35 dB of acoustic noise reduction around 600 Hz, ranging from 33 and 36 dB across individuals. The corresponding reduction in the broadband RMS noise level was 14 dB.

A typical adaptive run recorded from listener 4 is illustrated in Figure 2A. The events in the track include the initial recording scan and the four training scans, followed by 38 stimulus trials. After the initial rapid training phase, the residual sound at the error microphone had reduced from 91 dB SPL down to 54 dB SPL — a drop of 37 dB. For every adaptive run, the training phase provided at least 30 dB of acoustic cancellation. A substantial benefit was reached over the first one minute or so of the adaptive run, followed by a period of further slow-rate optimization of the ANC system. This provided an additional 5 dB or so of acoustic benefit. However, the acoustic reduction was not absolutely stable over time and whether the better ear, acoustically speaking, was the left or the right varied from trial to trial.

*** Insert Table I and Figure 2 about here ***

Figure 3 shows the three examples of the masker envelopes taken from recordings at the error microphone. The top panel shows one of the non-cancelled maskers, illustrating the periodicity of peaks and dips caused by the sequential acquisition of the slices in the scan. The remaining panels show two of the cancelled maskers. The ANC system has reduced the level by 30-40 dB, but there is a noteworthy moment-to-moment variation in its performance. The middle panel shows a general improvement across the masker, but the system has not preserved the periodicity evident in the non-cancelled masker, whereas the bottom panel shows slightly better performance as well as an overall periodicity.

*** Insert Figure 3 about here ***

The amount of cancellation was on average 4 dB greater in the better ear. There was also a significant influence of gender on the final level of the cancelled masker in either ear (F[319,1] = 111.1, p<0.001). The level of the residual noise both before and after cancellation for the two male listeners (1 and 2, mean) was higher than for the two female listeners (3 and 4) and we speculate that the effect of head size may be important here. When the male listeners were wearing the headset, it was a tight fit to position them inside the headcoil and there was no available space between them and the coil itself. Although the sides of the headset were padded with two thin layers of foam, it is likely that this contact provided an additional route for the conduction of vibrated energy from the MR scanner. In contrast, the two female listeners had smaller heads and so there was no coupling between the headset and the headcoil.

The data presented in Table I also enable comparison between the acoustic levels of the masker and the target, measured at the same four reversal points in the adaptive run. On average, the level of the target was significantly lower than the masker in the better ear (paired t[159]=-3.5, p=0.001), but only by 1.5 dB. Table I reveals an interesting between-subject difference. Only listeners 2 and 4 were able to detect the 600-Hz target at a level that was reliably below that of the masker in the better ear (p<0.0125, Bonferroni adjusted).

2. Perceptual benefits of noise reduction

For each listener, the perceptual benefit of noise reduction was computed by subtracting the mean threshold during the ten adaptive runs in the non-cancelled condition from each of their thresholds measured in the cancelled condition. Across the four listeners, the average detection

threshold for the 600-Hz target masked by the non-cancelled masker was 74 dB SPL. With cancellation, detection thresholds reached 54 dB SPL, an improvement of 20 dB. Table II reports the size of the perceptual benefit of noise reduction for individual listeners. There was a significant difference in benefit between listeners ($F[39,1] = 5.4$, $p<0.01$), with post-hoc testing revealing that listener 4 gained significantly greater perceptual benefit from cancellation than listeners 1 and 2 ($p<0.05$).

** *Insert Table II about here* **

These results clearly demonstrate that the ANC system can provide useful amounts of both acoustic and perceptual noise reduction under true fMRI conditions. Our listeners gained about 35 dB of acoustic reduction and 20 dB of perceptual benefit. A perceptual benefit of 20 dB, through a reduction in the level of the EPI noise, would be valuable for future auditory fMRI experiments. Furthermore, allowing the filter coefficients of the ANC system to slowly adapt allowed this level of performance to be maintained throughout the adaptive run.

The perceptual benefit was nevertheless limited in some way as it remained smaller than the acoustic reduction. Some of the difference may be due to listeners taking advantage of the regular dips between the 'pings' of the non-cancelled masker (Figure 3, top panel), but not being able to take similar advantage from the irregular, and likely mostly unpredictable, envelopes of the cancelled masker (Figure 3, bottom panel). It is well known that signal-detection thresholds can depend upon the 'peakiness' of the envelope of the masker (e.g., by adding masker components in 'Schroeder phase' or 'sine-phase'; Kohlrausch and Sander, 1995). One study of some relevance to the present experiment is that of Gockel *et al.* (2002), who compared detection thresholds for noise signals masked by a 62.5 Hz-F_0 complex tone with either a repeated,

modulated envelope or with a random envelope. They found that the detection threshold in the former condition was between 5 and 20 dB lower than in the latter condition, depending on the masker level. These data would be consistent with the dip-listening suggestion here, though it should be noted that Gockel *et al.* used a broadband signal and masker. Moreover, inspection of Figure 3 shows that a signal level of about 70-75 dB (corresponding to the signal threshold in the non-cancelled condition) is similar to the level of the dips in the non-cancelled masker, while a signal level of about 45-60 dB (corresponding approximately the signal threshold in the cancelled condition) is less clearly related to the pattern in the envelopes of the cancelled maskers.

It is also possible that some of the difference between the acoustic and perceptual benefit can be attributed to the emergence of bone conduction as the primary transmission path of the EPI noise to the cochlea. The current results do not decide this issue, but it would nevertheless be worth future emphasis to reduce it (see Ravicz *et al.*, 2001).

IV. EXPERIMENT 2: BENEFITS FOR DETECTABILITY OF SOUND-RELATED BRAIN ACTIVITY MEASURED USING fMRI

Given that the ANC system significantly improves the detectability of a 600-Hz target that coincides with the main peak in the spectrum of the EPI noise, cancellation should also improve the statistical detection of frequency-dependent activation for a 600-Hz signal in an fMRI experiment. We expected to obtain an improved activation pattern with cancellation not only because the target is more easily perceived, but also because the effect of the background noise

in the baseline condition is reduced. The rationale for the latter effect is as follows. The most straightforward fMRI paradigm detects sound-related activity in a relative manner by subtracting brain images acquired during a baseline condition in which the sound is absent from those for which sound stimuli are presented. Given that the EPI noise itself is known to evoke significant auditory cortical activity (e.g., Bandettini et al., 1998; Hall et al., 2000; Ulmer et al., 1998), a high intensity of the background noise elevates the baseline of the auditory cortical response, thus reducing the dynamic range for detecting the additional activity evoked by the stimulus of interest.

Accordingly, Experiment 2 sought to demonstrate the neuroimaging benefit of the ANC system by measuring frequency-dependent activation in three listeners for signals presented at two intensities, with and without cancellation, in a same-different frequency-discrimination task. This paradigm addresses the most fundamental question in auditory fMRI, namely the sensitivity to detecting sound-related activity. The attenuated condition presents an interesting case in which cancellation exerts the greatest effect on stimulus audibility. The target should be easy to hear when the ANC system is cancelling the masker noise, since it will be well above the detection threshold, but the target should be inaudible when the ANC system is not operational.

A. Methods

1. Stimuli

The stimuli were constructed in the same manner as for Experiment 1, except that they had center frequencies of 570, 600, or 630 Hz, and a duration of 450 ms. Stimuli were concatenated

to form pairs, separated by a 50-ms silent gap. In total, there were 60 same-frequency pairs (20 of each combination of 570/570 Hz, 600/600 Hz and 630/630 Hz) and 60 different-frequency pairs (30 of each combination of 570/600 Hz and 600/630 Hz). The order of the two stimuli in the different-frequency pairs was counterbalanced. The total RMS power was equated across stimulus pairs. Each condition contained a sequence of eight stimulus pairs, four same-frequency and four different-frequency trials, presented in a randomized order.

Six conditions were formed by crossing three signal levels ("94 dB", "attenuated" and "no signal") with two ANC states ("non-cancelled" and "cancelled"). The 94-dB condition represents the maximum possible presentation level of the 600-Hz signal and the stimuli at this level were clearly audible in both cancelled and non-cancelled conditions. The attenuated condition was equivalent to 3 dB *below* that individual's mean 600-Hz detection threshold in the *non*-cancelled ANC condition (as measured in Experiment 1). The audibility of the attenuated stimuli was therefore dependent on the ANC state. The physical levels of the stimuli and the resulting sensation levels for each listener are reported in Table III. In the no-signal condition, the signal level was not present. This condition served as the baseline comparison in a typical subtraction analysis of the fMRI data.

*** Insert Table III about here ***

A 'trial' comprised a stimulus pair that was presented at the same time as the MR scanner acquired an 18-slice set of brain images. These scans, and the characteristics of their EPI noises, were the same as those used in Experiment 1 (see Figure 1). However, in the present experiment, a conventional MR pulse sequence was used, i.e., sets of brain images were acquired repeatedly (at 2.7 s intervals) and there were no dummy scans. The onset of the stimulus pair always

occurred 0.05 s after the onset of the masker, and again stimulus pairs were presented at a rate of one every 2.7 s.

2. Procedures

Listeners were required to decide whether the two members of each pair were the same or different and to respond accordingly by pressing one of two buttons with the index finger of their right or left hand. The quiet interval of 1.55 s between trials was sufficient to allow this response. Listeners were instructed to guess when their response was uncertain and also to guess when the stimulus pair was inaudible.

The presentation order was carefully chosen to ensure equal distribution of the six conditions across four, 9-minute experimental runs. The conditions cycled in a pseudo-random order, such that there were never more than 64 stimulus pairs between repetitions of the same condition. In addition, throughout the experiment, the ANC system cycled between cancelled and non-cancelled conditions after every 24 stimulus pairs (i.e. after every three conditions). Each of the six conditions occurred 16 times throughout the whole experiment, giving a total dataset of 6x128 (N=768) trials and the same number of brain scans.

The fMRI procedure positioned the brain slices in the same orientation as in Experiment 1. To optimise coverage of the brain within a slice, the in-plane field of view was chosen to be 192 mm with a 64x64 matrix, corresponding to an image resolution (i.e. voxel size) of 3x3x3 mm. At the end of the four experimental runs, we also acquired a 1x1x1 mm resolution anatomical scan of the whole brain to enable precise anatomical localization of auditory cortical activity. The anatomical scan took 4½ minutes.

3. Listeners

Three right-handed listeners participated in Experiment 2. They were listeners 1, 2 and 4 from Experiment 1.

4. fMRI analysis

A standard fMRI analysis was conducted separately for each listener using Statistical Parametric Mapping v2 (SPM2) software (www.fil.ion.ucl.ac.uk/spm) running on MATLAB v6.5.0. A number of spatial processing steps were initially carried out on the EPI scans. To minimise the minor head movements that typically occur over the scanning period, all of the scans in the four runs were realigned using the last scan of the first run as the reference. Within runs, the adjustment for movement was never more than 3 mm and 1º, indicating that the listeners had complied with the instructions to remain as still as possible. To interpret the anatomical location of sound-evoked activation using a standard atlas of the brain and to compare across different brains, the individual brain images were next transformed into the brain space defined by the template of the Montreal Neurological Institute. The aim of this procedure is to match the cortical volume of each individual brain to the 'standard' brain. An automated algorithm minimised the sum of squares difference between the anatomical scan and the template, first by determining the optimum 12-parameter affine transformation, followed by estimating nonlinear deformations, whereby the deformations are defined by a linear combination of three dimensional discrete cosine transform basis functions. These deformation fields were then applied to the EPI scans, up-sampling the transformed images to a voxel resolution of 2x2x2

mm. The final spatial processing step involved smoothing these scans by a Gaussian kernel of 4 mm full-width-at-half-maximum to improve the signal-to-noise ratio. The spatially transformed images of each participant provided the input to the analysis of the sound-evoked brain activation. The analysis used the General Linear Model that is implemented in SPM2. The model partitions the observed response according to a sum of weighted variables (defining the experimental conditions). We used a model that specified the chronological sequence of events using 12 experimental variables for each run; six comprised the main sustained response associated with each experimental condition and six others were included to capture any differences in the latency of the peak response from the canonical form of the haemodynamic response function by using its temporal derivative. A final set of four variables accounted for any large between-run differences in the mean image intensity. Low-frequency artefacts in the time series were dealt with by applying a high-pass filter with a cut-off of 0.003 Hz. After model estimation, statistical contrasts between conditions were specified by linear combinations of these variables and the significance of each contrast was determined relative to the scan-to-scan residual variability. Four major contrasts were specified. Each contrast was a pair-wise comparison between a particular sound condition (namely; the (i) cancelled 94-dB, (ii) cancelled attenuated, (iii) non-cancelled 94-dB, or (iv) non-cancelled attenuated) and its corresponding 'no signal' baseline condition, and with a probability criterion for statistical significance of 0.001. For listeners 1 and 2, the variables for the main sustained response were sufficient to capture sound-evoked brain activity, whereas for listener 4 it was necessary to add in the variables for additional shifts in the onset and offset of the evoked response.

In addition to visually inspecting the activation patterns, the extent of sound-evoked activity was quantified using two auditory regions of interest that were defined using the outer-most borders from two standardized templates. The primary auditory cortex on Heschl's gyrus was defined using criteria from Morosan *et al.* (2001), while the posterior non-primary auditory cortex on the planum temporale was defined using criteria from Westbury *et al.* (1999). The latter region excluded any voxels that overlapped with the primary region to ensure that voxels in the two regions were mutually exclusive.

B. Results

1. Behavioural performance

All three listeners reliably judged (d' > 1.2) whether the stimulus pair were same or different in all the three conditions in which the stimuli were clearly detectable (c.f. Reed and Bilger, 1973), including the attenuated signals when the masker was reduced by cancellation (Figure 4). In contrast, performance was poor (d' < 0.2) for the same attenuated signals presented without any cancellation.

*** Insert Figure 4 about here ***

2. Sound-evoked brain activity

Brain activity was examined by overlaying voxels that showed a significant pair-wise effect onto the listener's anatomical scan. Suprathreshold activation occurred in auditory cortical regions, but not in subcortical structures, such as the medial geniculate body. For all three listeners,

activity encompassed parts of the primary auditory cortex on Heschl's gyrus and the posterior non-primary auditory cortex on planum temporale, although for listener 1 activation was more restricted to lateral regions. This is illustrated in Figure 5, which shows a benefit of cancellation in the attenuated condition for all three listeners. The extent of auditory cortical activity is quantified in Figure 6. Although the extent of sound-evoked activity was reduced for stimuli presented at a lower level compared to the 94-dB condition, a significant advantage for using ANC emerged for the non-primary auditory region ($t[5] = 4.5$, $p<0.01$). There was also a trend towards an advantage in primary auditory cortex, but this improvement did not reach significance ($p=0.21$).

*** Insert Figures 5 and 6 about here ***

In summary, Experiment 2 demonstrated that for individual listeners, ANC offers significant advantages for the fMRI detection of sound-related activity at low intensities. The extent of activity did *not*, however, simply increase monotonically with the sensation level of the stimulus. This result is due to the fact that sound-evoked activation is determined by subtracting the baseline from the stimulus conditions and so it is influenced by the level of the EPI noise, as well as the level of the stimulus (see also Hall *et al.,* 2001; Hart *et al.*, 2003; Scott *et al.*, 2004). The particular benefit provided by ANC for detecting sound-related activation in non-primary regions may reflect the decreased listening effort required to segregate stimuli from the background noise (Griffiths and Warren, 2002).

III. EXPERIMENT 3: EFFECTS OF CANCELLATION AT OTHER FREQUENCIES OF THE EPI NOISE

Experiment 3 extended the main hypothesis explored in Experiment 1 by investigating the acoustical and perceptual benefits of cancellation for target stimuli presented at frequencies away from the prominent peak of energy in the EPI noise. In this experiment, we measured detection thresholds for narrow bands of noise whose center frequencies corresponded to other prominent low-frequency peaks in the spectrum of the EPI noise — namely 150 and 300 Hz — and which were attenuated by at least 10 dB during the cancellation. We also examined the effects for an additional frequency (450 Hz) that corresponds to a 'dip' in the spectrum of the EPI noise.

A secondary goal was to explore the effect of fixing the coefficients of the adaptive filter after the initial training period and to document the stability of the noise reduction over time. Fixing these coefficients effectively 'freezes' the cancellation signal. This procedure was initially deemed a reasonable choice for ANC implementation because both the gradient coils that generate the EPI noise and the trigger pulse that synchronises the DSP contain a clock that is controlled by a crystal oscillator, and are therefore highly stable over successive repetitions (Chambers *et al.*, 2007). For example, Chambers *et al.* (2007) reported that time variations ($< \pm 0.1$ μs) between a trigger being received and the DSP system starting a conversion were insufficient to affect cancellation performance. Documenting the levels of the masker noise and the target over time enabled us to quantify the performance of the 'fixed' ANC system.

A. Methods

1. Stimuli

The targets were diotic narrowband noises centered on 150, 300, 450 and 600 Hz. The stimuli were created in the same way as for Experiment 1. They had 50-Hz bandwidth and 500-ms duration (including 50-ms raised-cosine ramps applied to the onset and offset) and were sampled at a rate of 44,100 Hz. The presentation level was varied according to the adaptive procedure by applying a digital attenuation before digital-to-analogue conversion. Again for the 600 Hz target, a 0-dB attenuation corresponded to a level of 94 dB SPL. For the 150-, 300- and 450-Hz targets, 0-dB attenuation corresponded to levels of 94, 95, and 94 dB SPL respectively. Details of the acoustic structure of the masking noise are reported in Experiment 1 (section II.A.2). As before, the target was presented randomly within one of the two maskers and its onset always occurred 0.40 s after the onset of the masker.

2. Procedures

We used the same bespoke EPI pulse sequence, as described for Experiment 1 (section II.A.2), to generate a noise sequence that would permit us to conduct a two-interval psychophysical paradigm. Detection thresholds for the targets were measured using 2I-2AFC trials allied to a three-down, one-up adaptive procedure to track 79% correct. The listener was required to judge whether the target appeared in the first or second interval, responding appropriately by pressing one of two buttons with the index finger of either their left or right hand. As in Experiment 1, listeners completed an earlier practice-and-familiarization session detecting the 600-Hz target in the plastic-and-wood replica of an MR scanner.

The procedure for initializing the ANC system was as described before (section II.A.1). The only exception was that the coefficients of the adaptive filter (and hence, the characteristics of the cancellation signal) were fixed after the training scans. On the first two practice trials, the initial attenuation levels of the target were 15 dB in the cancelled condition and 5 dB in the non-cancelled condition, hence the target always started off being clearly audible. Thereafter, the level changed according to performance in 5-dB steps for the first two reversals, and 2-dB steps for subsequent reversals. For the first two listeners (5 and 6), each adaptive run ended after four 2-dB reversals. Three runs were repeated for each condition and the mean detection threshold was taken as the average of the levels at the last four reversal points across runs. Generally, each run contained 30-40 trials and lasted about 3½ minutes. Although a total time of 21 minutes was required to obtain the cancelled and non-cancelled thresholds for one target frequency, the significant amount of setting-up time in-between runs made these sessions very long and tiring. Consequently, the final two listeners (7 and 8) completed just two runs for each condition, using instead a termination rule of eight 2-dB reversals. The mean detection threshold was taken as the average of the levels at the last eight reversal points across runs. Using this protocol, each adaptive run contained 50-60 trials and lasted about 5½ minutes. Although the total acquisition time remained about the same as before, the sessions were considerably shortened by the reduction in setting-up time. During each adaptive run, the residual sound was monitored by the error microphone under the left and right ear defenders and recorded onto a DAT recorder (Sony DTC-690) for off-line evaluation of the acoustic cancellation over time. Individual measurements were completed in two 1½-hour testing sessions, the order of conditions being counterbalanced across sessions and listeners.

3. Listeners

Four listeners took part in Experiment 3, one male (5) and three females (6-8). Listeners were recruited from the undergraduate population at Nottingham University and had a median age of 21.5 years (range 20-22). All listeners had normal hearing as assessed using standard pure-tone audiometry (< 20 dB HL over octave frequencies between 250 and 8000 Hz).

B. Analysis

The DAT recordings were analyzed to determine the sound level at the error microphone over the course of each adaptive run using the procedures reported in Experiment I (section II.B.). Again, we were interested only in the amount of cancellation at the frequency corresponding to the target and so the digitized recordings were first filtered to remove sound energy that was remote from the frequency of interest. We chose bandwidths of 60, 80 and 100 Hz as convenient approximations to the ERBs for 300, 450 and 600 Hz respectively (Glasberg and Moore, 1990). Filtering the digitized recordings for the 150-Hz condition, using a filter bandwidth of 40 Hz, did not yield meaningful estimates because the level of the cancelled masker approached that of the ambient noise. This was continually present and primarily reflects the air-conditioning in the room.

As before, the perceptual benefits of noise reduction for each individual were computed by subtracting the mean detection threshold in the non-cancelled condition from the mean detection threshold in the cancelled condition.

C. Results

1. Acoustic noise reduction during the cancelled condition

The results for target frequencies of 300, 450 and 600 Hz demonstrated that the ANC system operated effectively across a range of low frequencies and attenuated both peaks *and* dips in the acoustic spectrum. Table IV shows the non-cancelled and cancelled masker levels for each center frequency, computed from the recording made at the microphone under the earcups during the tests. On average, the ANC system provided similar amounts of acoustic noise reduction at 300 and 600 Hz (22 dB and 21 dB respectively), but somewhat less noise reduction at 450 Hz (10 dB) [4]. Once again, there was some asymmetry in the degree of cancellation across the left and right channels, with the residual masker level being about 4 dB lower in the better ear.

** *Insert Table IV about here* **

We quantified the effect of fixing the adaptive-filter coefficients by comparing the performance of the ANC system at 600 Hz in Experiment 3 with its corresponding performance in Experiment 1, in which the coefficients were allowed to adapt slowly over the run. We found that the acoustic noise reduction was diminished in Experiment 3 compared to Experiment 1 (21 dB versus 35 dB). This is due to the residual level of the masker being generally much higher in Experiment 3 (72 dB versus 58 dB SPL). An illustrative example of the performance of the ANC system over the course of an adaptive run is shown in Figure 2B (recorded from listener 6). Despite an initial noise reduction of over 30 dB the amount of noise reduction began to diminish immediately after the second training scan (the moment at which the adaptive filter coefficients were fixed). Thereafter, there was a steady deterioration over time, especially in the right ear.

2. Perceptual benefits of noise reduction

The mean perceptual benefits are reported in Table II for the four different center frequencies. As expected, the perceptual benefit was greatest at 600 Hz, the peak energy of the EPI noise, although reliable improvements were obtained at all other frequencies. On average, listeners were able to reliably detect the targets at a level below the level of the better ear masker (paired $t[155]=-23.1$, $p<0.001$) (see Table IV). This pattern had been observed previously in two of the listeners that participated in Experiment 1.

Again, we quantified the effect of fixing the adaptive-filter coefficients by comparing the individual target detection thresholds at 600 Hz across Experiments 1 and 3 (see Table II). Although the non-cancelled thresholds were broadly similar, the cancelled thresholds were always higher in Experiment 3, in which the cancellation signal was fixed. This result is consistent with the view that listening performance was in the cancelled condition was compromised by the poorer acoustic performance of the ANC system due to the fixing of the adaptive-filter coefficients

IV. GENERAL DISCUSSION

Taken together, these experiments demonstrate that active noise cancellation can be achieved in the electromagnetically hostile and physically compact neuroimaging environment. In practice, the ANC system does not involve a significant increase in scanning time and can be used with minimal training. Although the present paper reports the levels of acoustic noise

reduction over time, these detailed measurements are not necessary under normal operating conditions. More routine on-line monitoring of the residual sound at the listener's ears is provided by two level meters on the front of the DSP system. Consequently, routine use of the ANC system does not limit the feasibility of scanning large numbers participants (N=10-20). We therefore recommend that active noise reduction of the EPI noise should be implemented whenever possible during auditory fMRI research.

As well as improving the audibility of sounds presented to participants, ANC increased the sound-evoked fMRI activity for stimuli presented at low levels. The noise reduction was obtained across a range of low frequencies corresponding to both peaks and dips in the spectrum of the EPI noise and reduction was greatest when the ANC system continually adjusted the cancellation signal to maintain optimal performance over time.

One of the main advantages for applications of auditory fMRI that is offered by active noise cancellation is that it enables experimenters to present sound stimuli at much lower levels than has been possible hitherto during functional neuroimaging studies. Typically in fMRI studies, acoustic signals are delivered at moderate to high sound levels (see Hall, 2005 for a review). For example, for mapping frequency-dependent responses across auditory cortex, Talavage *et al.* (2000) presented sounds at 42-50 dB above the threshold measured with the participant in the scanner. Although fMRI research has made important advances in characterising the functional anatomy of human tonotopy, especially in the primary auditory region on the medial two-thirds of Heschl's gyrus, it is important to remember that the response pattern of a given neuron is not rigidly fixed, but can change as a function of other acoustic features, including level. Thus, it may not be valid to directly compare these human fMRI results

with data from microelectrode recordings in animals that localize the best frequency response near threshold, because the cortical area responsive to a suprathreshold stimulus cannot be predicted from the threshold map (Phillips *et al.*, 1994). ANC offers the possibility to re-examine issues such as human tonotopy using sounds presented at a level up to 20 dB below the non-cancelled equivalent.

The application of ANC also raises an opportunity to address other questions that have previously been confounded by the excessive EPI noise levels. One example concerns measures of the neural correlates of speech intelligibility. Davis and Johnsrude (2003) have shown a clear association between the signal-to-noise ratio and speech-related activation in many brain areas including the superior and middle temporal gyri, left inferior frontal gyrus and left hippocampus. These brain areas respond to the intelligibility of the speech signal. Clearly, a reduction in background noise level by 20 dB should markedly improve speech intelligibility and hence, the resulting pattern of activity. A second example concerns measures of the neural basis of temporal coding for complex tones (Hall *et al.*, 2006). Coding of periodicity pitch can be examined using harmonic complex tones in which the first few harmonic components are missing from the spectrum. However, nonlinear interactions on the basilar membrane can reconstitute the lower part of the harmonic series as a series of distortion products. Psychophysical evidence demonstrates that, although the components of the distortion spectrum are reduced in level by 10 dB and more compared to the primaries, many exceed the hearing threshold (Pressnitzer and Patterson, 2001). Furthermore, the higher the presentation level of the complex tone, the greater the number of distortion components above threshold. Clearly, a reduction in the presentation

level by up to 20 dB, through the use of ANC, could reduce the contribution of such confounds in fMRI studies of pitch coding.

Have we reached the limits on the amount of attenuation that can be achieved? Our observations concur with previous investigations of passive and active noise reduction measured at low frequencies (< 1 kHz). Ravicz *et al.* (2001) evaluated different combinations of earplugs, ear defenders and/or a sound-absorbing helmet and concluded that (at least for frequencies < 500 Hz) the ear canal was the dominant conduction route for every device configuration tested. Using an ANC system in a laboratory setting to measure detection thresholds for a tone masked by white noise, Chambers *et al.* (2001) had found that the detection thresholds were improved by a substantial amount, but not quite as much as the acoustic noise reduction at frequencies below 1 kHz. For example, at 500 Hz, the average perceptual and acoustic benefits were 17 and 23 dB, respectively. The maximum perceptual benefit that can be gained by using methods that treat only sound conduction through the head (e.g. ear defenders combined with ANC) is determined by whether the dominant transmission path of sound to the cochlea is airborne or bone-conducted. It is likely that accurate estimates of the upper bound of noise reduction for vibrations that are mechanically coupled to the head (as in this case when listeners are lying in the head coil) would be difficult to ascertain, and indeed have not been attempted to the best of our knowledge. Standard laboratory measurements of free-field bone conduction are often made using one or more loudspeakers at some distance from the listener (e.g. Berger, 1983; Berger and Kerivan, 1983) and so may not be directly applicable to the MR environment, where the listener is lying *inside* the source of the noise and, moreover, the source is physically large and probably cannot be assumed to be a point source. The question remains as to whether further

improvements to the performance of the ANC system might confer even greater perceptual benefit. For instance, placing the error microphone nearer to the listener's ear drum rather than just under the ear defender may provide additional improvements in the efficacy of the cancellation signal.

ACKNOWLEDGEMENTS

Dr Yuvi Kahana at Optoacoustics Ltd, Or Yehuda, Israel, kindly provided the optical microphones. Ms Kimberley Quinn helped to test participants in Experiment 3 and also contributed to a pilot version of Experiment 2 as part of her final-year, BSc Physics project at Nottingham Trent University. We are also grateful to Brian Moore (Associate Editor) and two anonymous reviewers for their insightful comments on an earlier version of the manuscript.

FOOTNOTE

[1] The term 'noise' is commonly used in the imaging literature to characterise the distraction and annoyance of the sound produced by the scanner, and also to emphasise the desirability of attenuating, or removing, it. Perceptually, however, the acoustic signal generated by the scanner during EPI sounds more like a sharp, intense 'ping', with a clear tonal character, rising above a broadband spectrum of noise that is at a much lower intensity.

[2] The difference between bone and air conduction was measured by Berger (1983) using the 'real-ear attenuation at threshold' (REAT) method (e.g. Berger and Kerivan, 1983).

[3] This is the only period in which the scanner room is quiet. Although the scanner room itself was not sound treated, at 57 dB(A), the background level was well below the level of the EPI noise.

[4] Although Figure 1 indicates that there is no acoustic reduction at 450 Hz, our report of 10 dB of acoustic reduction is based on analysis using a narrowband filtered signal, with cut-offs at 410 and 490 Hz.

REFERENCES

Amaro, E. Jr., Williams, S. C. R., Shergill, S. S., Fu, C. H. Y., MacSweeney, M., Picchioni, M. M., Brammer, M. J., and McGuire, P. K. (**2002**). "Acoustic noise and functional magnetic

resonance imaging: Current strategies and future prospects," J. Magn. Reson. Imaging **16,** 497–510.

Bandettini, P. A., Jesmanowicz, A., Kylen, J. Van, Birn, R. M., and Hyde, J. S. (**1998**). "Functional MRI of brain activation induced by scanner acoustic noise," Mag. Reson. Med. **39**, 410–416.

Berger, E. H. (**1983**). "Laboratory attenuation of earmuffs and earplugs both singly and in combination," Am. Ind. Hyg. Assoc. **44**, 321-329.

Berger, E. H. and Kerivan, J. E. (**1983**). "Influence of physiological noise and the occlusion effect on the measurement of real-ear attenuation at threshold," J. Acoust. Soc. Am. **74**, 81-94.

Boyle, Y., Bentley, D.E., Watson A., and Jones, A. K. P. (**2006**). "Acoustic noise in functional magnetic resonance imaging reduces pain unpleasantness ratings," NeuroImage **31**, 1278-1283.

Chambers, J., Bullock, D., Kahana, Y., Kots A., and Palmer, A. R. (**2007**)."Developments in active noise control sound systems for magnetic resonance imaging," Appl. Acoustics **68**, 281-295.

Chambers, J., Akeroyd, M. A., Summerfield, A. Q., and Palmer, A. R. (**2001**). "Active control of the volume acquisition noise in functional magnetic resonance imaging: method and psychoacoustical evaluation," J. Acoust. Soc. Am. **110**, 3041-3054.

Chapman, B. L. W., Haywood, B., and Mansfield, P. (**2003**). "Optimized gradient pulse for use with EPI employing active acoustic control," Mag. Reson. Med. **50**, 931-935.

Cho, Z. H., Chung, S. C., Lim, D. W., and Wong, E. K. (**1998**). "Effects of the acoustic noise of the gradient systems on fMRI: a study on auditory, motor, and visual cortices," Mag. Reson. Med. **39**, 331–335.

Davis, M. H., and Johnsrude, I. S. (**2003**). "Hierarchical processing in spoken language comprehension," J. Neurosci. **23**(8), 3423–3431.

Elliott, M. R., Bowtell R. W., and Morris, P. G. (**1999**). "The effect of scanner sound in visual, motor, and auditory functional MRI," Mag. Reson. Med. **41**, 1230–1235.

Foster, J. R., Hall, D. A., Summerfield, A. Q. Palmer, A. R., and Bowtell, R. W. (**2000**). "Sound-level measurements and calculations of safe noise dosage during EPI at 3T," J. Mag. Reson. Imaging **12**, 157-163.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Gockel, H., Moore, B.C.J., and Patterson, R.D. (**2002**). "Asymmetry of masking between complex tones and noise: The role of temporal structure and peripheral compression," J. Acoust. Soc. Am. **111**, 2759-2769.

Granet, R., and Gelber, L. (**1990**). "Claustrophobia during MR imaging," New Jersey Med. **87**, 479–482.

Griffiths, T. D., and Warren J. D. (**2002**). The planum temporale as a computational hub. Trends Neurosci. **25**, 348-353.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R. Summerfield, A. Q., Elliott, M. R., Gurney, E., and Bowtell, R. W. (**1999**). "Sparse temporal sampling in auditory fMRI," Hum. Brain Mapp. **7**, 213-223.

Hall, D. A., Summerfield, A. Q., Gonçalves, M. S., Foster, J. R., Palmer, A. R., and Bowtell, R. W. (**2000**). "Time-course of the auditory BOLD response to scanner noise," Mag. Reson. Med. **43**, 601-606.

Hall, D. A., Haggard, M. P., Summerfield, A. Q., Akeroyd, M. A., Palmer, A. R., and Bowtell, R. W. (**2001**). "FMRI measurements of sound-level encoding in the absence of background scanner noise," J. Acoust. Soc. Am. **109**, 1559-1570.

Hall, D. A., Hart, H. C., and Johnsrude, I. S. (**2003**). "Relationships between human auditory cortical structure and function," Audiol. Neurootol. **8**, 1-18.

Hall, D. A. (**2005**). Representations of spectral coding in the human brain. In "Auditory Spectral Processing. International Review of Neurobiology, Volume 70" edited by M. S. Malmierca, D. R. F. Irvine (Elsevier, San Diego), pp 331-369.

Hall, D. A. (**2006**). "Haemodynamic-based imaging of central auditory function," in *Auditory Evoked Potentials: Basic Principles and Clinical Application*, edited by R. Burkard, M. Don and J. Eggermont (Lippincott, Williams and Wilkins, Baltimore), Chap. 2, pp. 22-40.

Hall, D. A., Edmondson-Jones, A. M., and Fridriksson, J. (**2006**). "Frequency and periodicity coding in human auditory cortex," Eur. J. Neurosci. **24**, 3601–3610.

Hall, J. W., and Fernandes, M. A. (**1984**). "The role of monaural frequency selectivity in binaural analysis," J. Acoust. Soc. Am. **76**, 435-439.

Hart, H. C., Hall, D. A., and 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," Hear. Res. **179**, 104-112.

Hennel, F., Girard, F., and Loenneker, T. (**1999**) "'Silent' MRI with soft gradient pulses," Mag. Reson. Med. **42**, 6-10.

Johnsrude, I.S., Giraud, A. L., and Frackowiak, R. S. J. (**2002**). "Functional imaging of the auditory system: The use of positron emission tomography," Audiol. Neurootol. **7**, 251–276.

Kohlrausch, A. and Sander, A. (**1995**). "Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets," J. Acoust. Soc. Am. **97**, 1817-1829.

Levitt, H. (**1971**). "Transformed up–down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

More, S. R., Lim, T. C., Li, M. F., Holland, C. K., Boyce, S. E., and Lee, J. H. (**2006**). "Acoustic noise characteristics of a 4 Tesla MRI scanner," J. Mag. Reson. Imaging **23**, 388-397.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (**2001**). "Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system," NeuroImage **13**, 684-701.

Nelson, P. A. and Elliott, S. J. (**1992**). "Active control of sound," Academic Press Ltd, London.

Phillips, D. P., Semple, M. N., Calford, M. B., and Kitzes, L. M. (**1994**). "Level-dependent representation of stimulus frequency in cat primary auditory cortex," Exp. Brain Res. **102**, 210-226.

Pressnitzer, D., and Patterson, R. D. (**2001**). "Distortion products and the perceived pitch of harmonic complex tones", in *Physiological and Psychophysical Bases of Auditory Function*, edited by A. J. M., Houtsma, A., Kohlrausch, V. F., Prijs, R., Schoonhoven (Shaker Publishing, Maastricht), pp. 97-104.

Quirk, M., Letendre, A., Ciottone, R., and Lingley, J. (**1989**). "Anxiety in patients undergoing MR imaging," Radiology **170**, 463–466.

Ravicz, M. E., Melcher, J. R., and Kiang, N. Y.-S. (**2000**). "Acoustic noise during functional magnetic resonance imaging," J. Acoust. Soc. Am. **108**, 1683-1696.

Ravicz, M. E., and Melcher, J. R. (**2001**) "Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: Examination of noise conduction through the ear canal, head, and body," J. Acoust. Soc. Am. **109**, 216-231.

Reed, C. M., and Bilger, R. C. (**1973**) "A comparative study of S/N and E/N_0," J. Acoust. Soc. Am. **53**, 1039-1045.

Schwarzbauer, C., Davis, M. H., Rodd, J. M., and Johnsrude, I. S. (**2006**). "Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI," NeuroImage **29**, 774-782.

Scott, S. K., Rosen, S., Wickham, L., and Wise, R. S. J. (**2004**). "A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception," J. Acoust. Soc. Am. **115**, 813–821.

Shah, N. J., Jancke, L., Grosse-Ruyken, M. L., and Muller-Gartner, H. W. (**1999**). "Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination," J. Magn. Reson. Imaging **9**, 19–25.

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

Ulmer, J. L. Biswal, B. B., Yetkin, F. Z., Mark, L. P., Mathews, V. P., Prost, R. W., Estkowski, L. D., McAuliffe, T. L., Haughton, V. M., and Daniels, D. L. (**1998**). "Cortical activation response to acoustic echo planar scanner noise," J. Comput. Assist. Tomogr. **22**, 111–119.

Westbury, C. F., Zatorre, R. J., and Evans, A. C. (**1999**). "Quantifying variability in the planum temporale: A probability map," Cereb. Cortex **9**, 392-405.

TABLES

Table I. Individual sound levels measured in dB SPL at the listeners' ears in Experiment 1. Values in parentheses represent the standard error of the measurement. The three different masker levels were computed from the DAT recording of the error microphone using a 100-Hz wide filter centered on 600-Hz. The masker levels reported for the cancelled condition correspond to those trials at the reversal points in the adaptive run. The individual threshold levels of the 600-Hz signal are provided for comparison with the masker levels in the better ear.

Listener	Non-cancelled	Cancelled			Acoustic noise reduction (better and worse ear mean)
	Masker (better and worse ear mean)	Masker in better ear	Masker in worse ear	Signal level	
1	94 (0.3)	57 *(0.3)*	62 *(0.7)*	59 *(0.7)*	37 *(0.5)*
2	94 *(0.7)*	59 *(0.3)*	62 *(0.6)*	55 *(0.7)*	37 *(0.4)*
3	92 *(0.6)*	53 *(0.3)*	59 *(0.4)*	53 *(0.7)*	38 *(0.4)*
4	90 *(0.2)*	53 *(0.3)*	57 *(0.5)*	50 *(0.6)*	37 *(0.3)*
mean	93	56	60	54	37

Table II. Masked perceptual thresholds in the cancelled and non-cancelled conditions for Experiments 1 and 3. Mean levels are reported in dB SPL, with standard errors in parentheses.

	Centre frequency	Listener	Non-cancelled threshold	Cancelled threshold	Perceptual benefit
Exp. 1					
	600 Hz	1	76 *(0.6)*	59 *(1.2)*	17 *(1.2)*
	"	2	73 *(1.0)*	55 *(1.3)*	18 *(1.3)*
	"	3	73 *(1.0)*	53 *(1.3)*	20 *(1.3)*
	"	4	73 *(1.7)*	50 *(1.1)*	23 *(1.1)*
Exp. 3					
	600 Hz	5	81 *(1.1)*	64 *(1.0)*	17 *(1.0)*
	"	6	76 *(0.7)*	65 *(1.2)*	11 *(1.2)*
	"	7	76 *(0.8)*	59 *(0.6)*	17 *(0.6)*
	"	8	81 *(0.8)*	66 *(0.8)*	15 *(0.8)*
Exp. 3					
	150 Hz	mean	69 *(0.5)*	65 *(0.4)*	4 *(0.6)*
	300 Hz	"	69 *(0.8)*	59 *(0.5)*	10 *(1.2)*
	450 Hz	"	64 *(1.1)*	54 *(0.6)*	10 *(1.4)*
	600 Hz	"	78 *(1.1)*	63 *(0.6)*	15 *(1.1)*

Table III. The physical levels of the stimuli and the resulting sensation levels, for each listener who participated in Experiment 2. The sensation levels for the 94-dB conditions are determined from the individual cancelled and non-cancelled thresholds reported in Experiment 1 (Table II).

Stimulus	ANC state	Sound level, dB SPL			Sensation level, dB		
		Listener			Listener		
		1	2	4	1	2	4
94 dB	non-cancelled	94	94	94	18	21	21
94 dB	cancelled	94	94	94	35	39	44
attenuated	non-cancelled	73	70	70	-3	-3	-3
attenuated	cancelled	73	70	70	14	15	20

Table IV. Mean sound levels measured in dB SPL at the listeners' ears across different frequencies of interest in Experiment 3. The masker levels reported for the cancelled condition correspond to those trials at the reversal points in the adaptive run. The mean threshold level of each target signal is provided for comparison with the masker levels in the better ear. Values in parentheses represent the standard error of the measurement.

Centre frequency	Recording scan (non-cancelled)	Cancelled			Acoustic noise reduction (better and worse ear mean)
	Masker (better and worse ear mean)	Masker in better ear	Masker in worse ear	Signal level	
150 Hz	-	-	-	51 *(0.4)*	-
300 Hz	91 *(0.8)*	66 *(0.5)*	72 *(0.6)*	59 *(0.5)*	22 *(1.0)*
450 Hz	75 *(0.9)*	62 *(0.6)*	68 *(0.2)*	54 *(0.6)*	10 *(2.3)*
600 Hz	94 *(0.4)*	70 *(0.7)*	75 *(0.9)*	63 *(0.6)*	21 *(2.0)*

COLLECTED FIGURE CAPTIONS

Figure 1. The average power spectrum of the scanner noise generated during functional brain imaging (EPI) using a Philips Intera 3 Tesla MR scanner. The black line (non-cancelled) indicates the spectrum of the noise recorded under normal operation, and shows a peak of power at the fundamental frequency of 600 Hz and lesser peaks at 150 and 300 Hz. The grey line (cancelled) indicates the spectrum when the ANC is operative. The power around the dominant peak is reduced by 35 dB. The inset (upper right) shows the time-domain waveform of a single burst of EPI noise, a sequence of 18 discrete 'pings' corresponding to a series of 18 consecutive slices through the brain. Each 'ping' lasts up to about 0.06 s, with the total duration of the burst being about 1.15 s.

Figure 2. Examples of the time course of acoustic noise reduction and its relationship to detection thresholds for a 600-Hz target. Diamond symbols denote the RMS level at the left ear and square symbols denote the RMS level measured at the error microphone at the right ear. The first scan corresponds to the recording of the noise reference and therefore indicates the non-cancelled level of the EPI noise under the ear defender, at just over 90 dB SPL. **A.** Example taken from listener 4 in Experiment 1. The level of the EPI noise rapidly reduces over the four training scans (grey infills) as the ANC system optimises the adaptive filter. In each pair of subsequent trials, the level in one interval is higher than in the other because it contains both the masker and the target signal. The black dots denote the actual presentation level of the 600-Hz target, the first two practice trials being at 79 dB SPL. The target level tracks the interval that

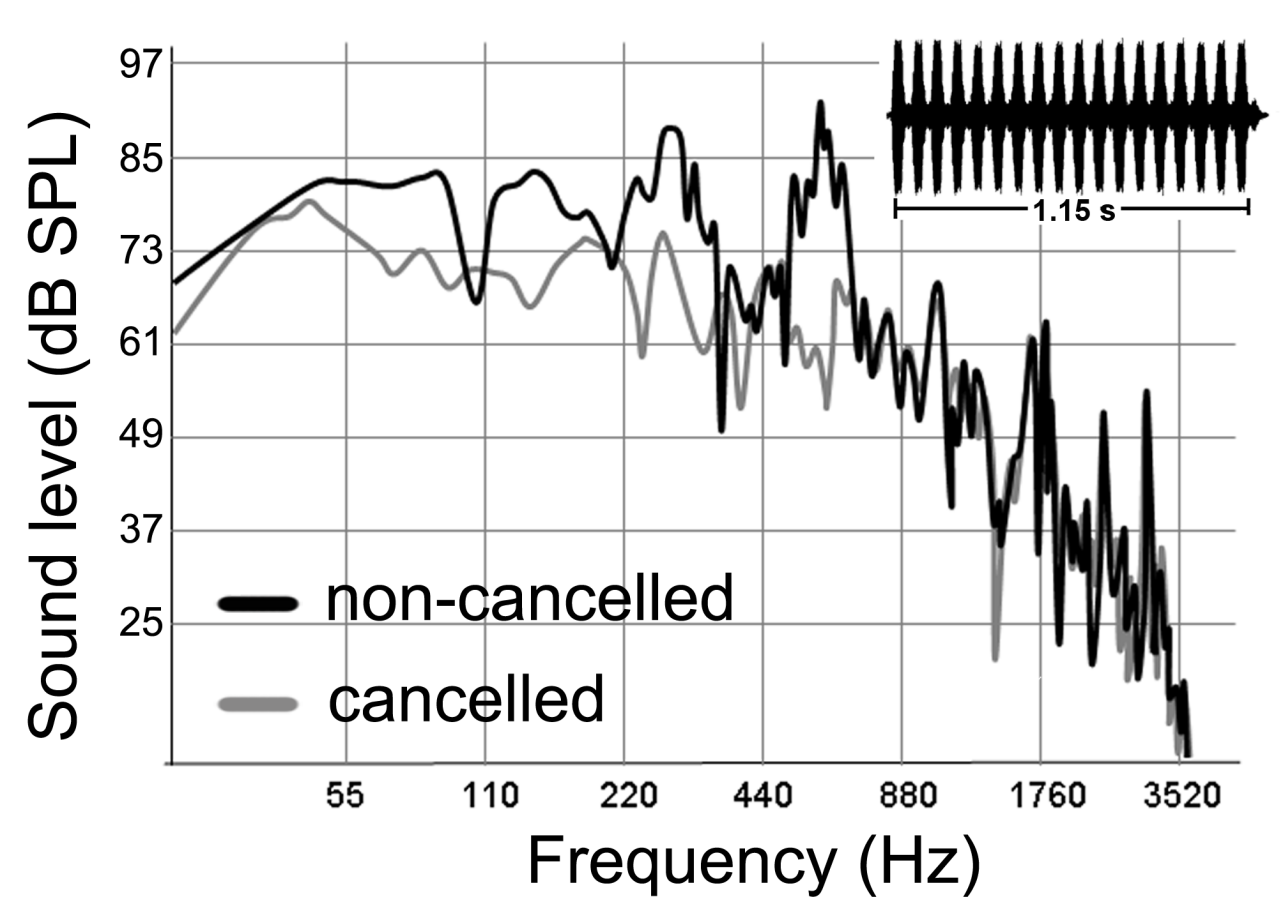
represents the combined level of the noise and target reasonably well. **B.** Example taken from listener 6 in Experiment 3. The level of the EPI noise rapidly reduces over the two training scans (grey infills), but the amount of noise reduction is generally less than Experiment 1 and the masker attenuation in the right ear can be seen to deteriorate by over 15 dB over the course of the adaptive run.

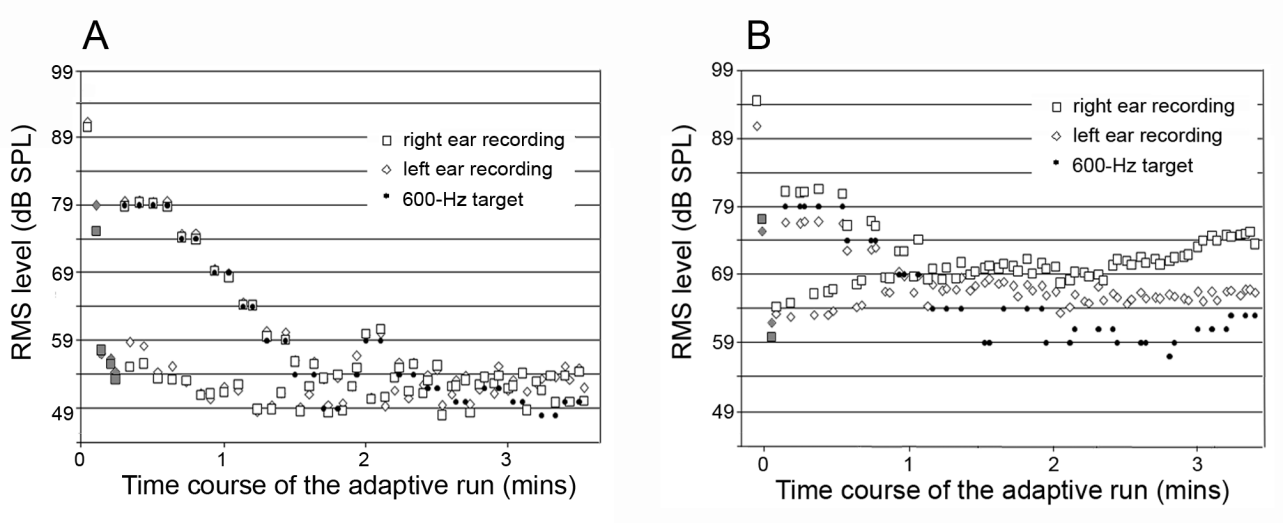
Figure 3. Envelopes of the recordings of three noise maskers taken from one adaptive track of listener 4. The top panel shows one of the non-cancelled maskers at the beginning of the run, and the other two panels show two separate cancelled maskers from later in the run. The envelopes were calculated using the Hilbert transform, using edited waveforms of 1500-ms duration that were approximately centered on the scan. Note that the waveforms were initially filtered to a 100-Hz wide band, centered on 600 Hz (see section II.B)].

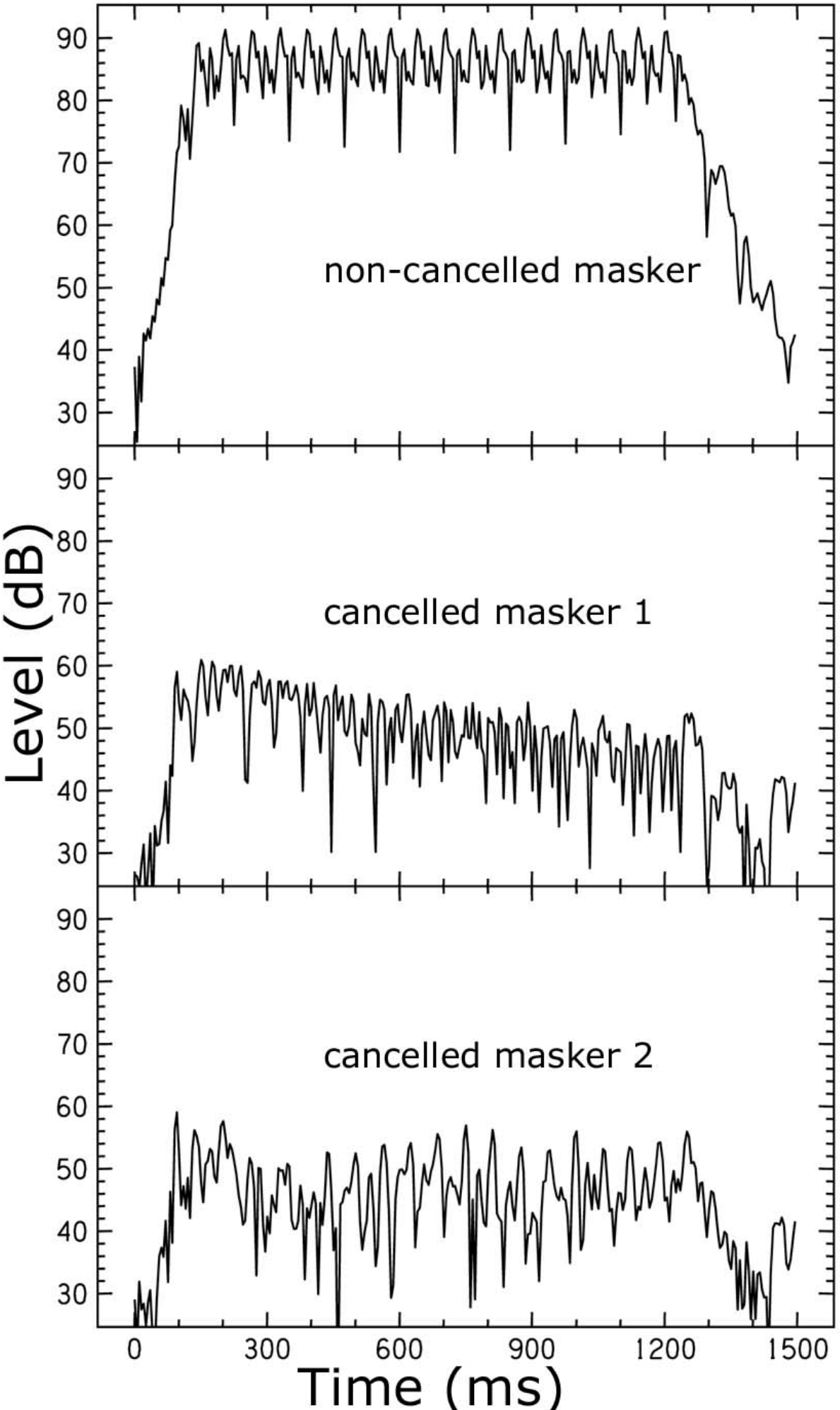
Figure 4. Same/different frequency discrimination performance for Experiment 2 measured using d'. Mean performance is shown by the height of the bar, whilst individual scores are represented by the ellipse (listener 1), the square (listener 2) and the circle (listener 4).

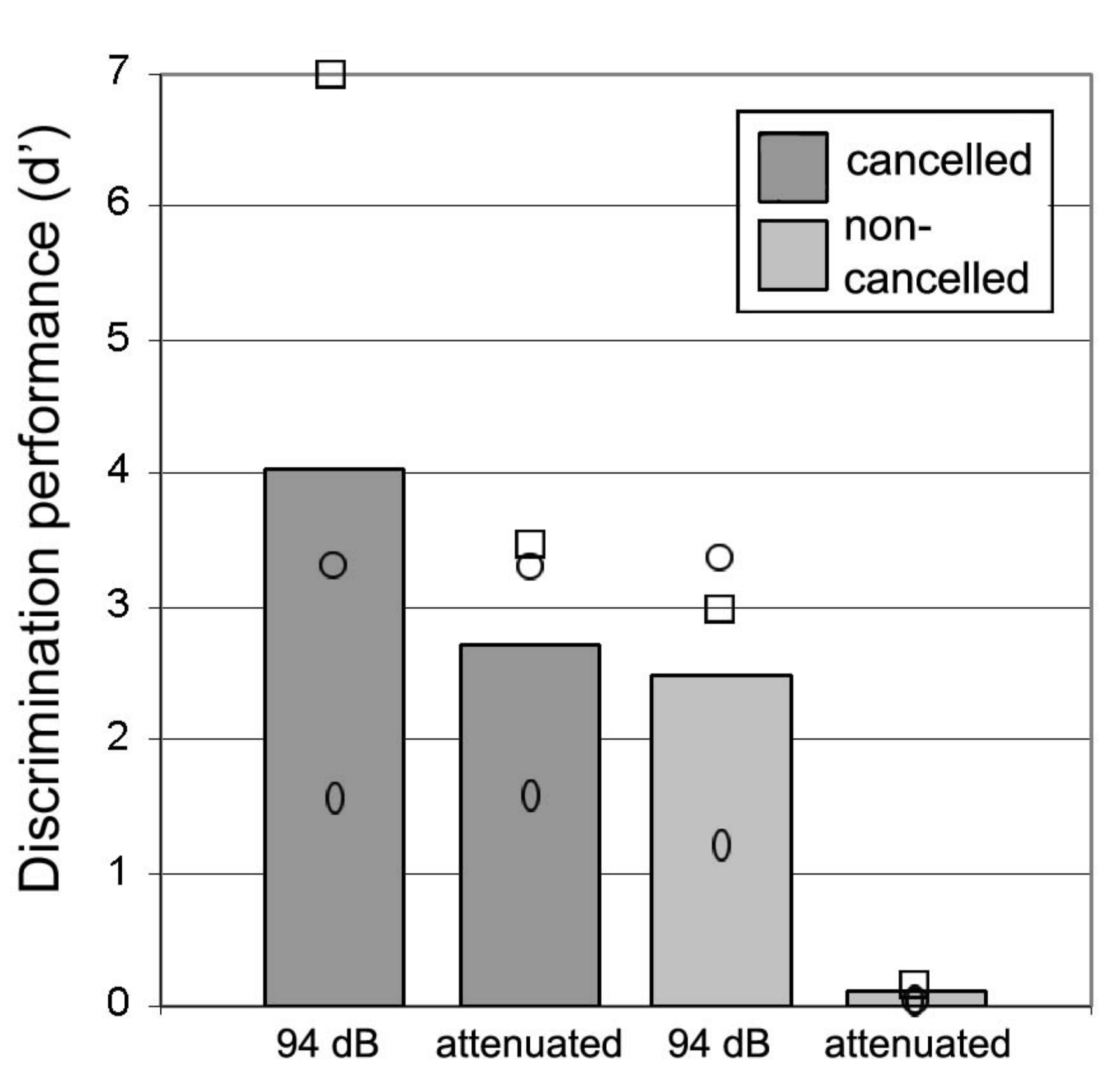
Figure 5. Sound-evoked brain activity for Experiment 2. Each row displays the results for an individual listener overlaid onto a slice that is oriented through auditory cortical regions on the upper surface of the superior temporal gyrus. Each column displays the pattern of significant brain activity, in white, for one of the four major contrasts (voxel-level threshold, $p<0.001$).

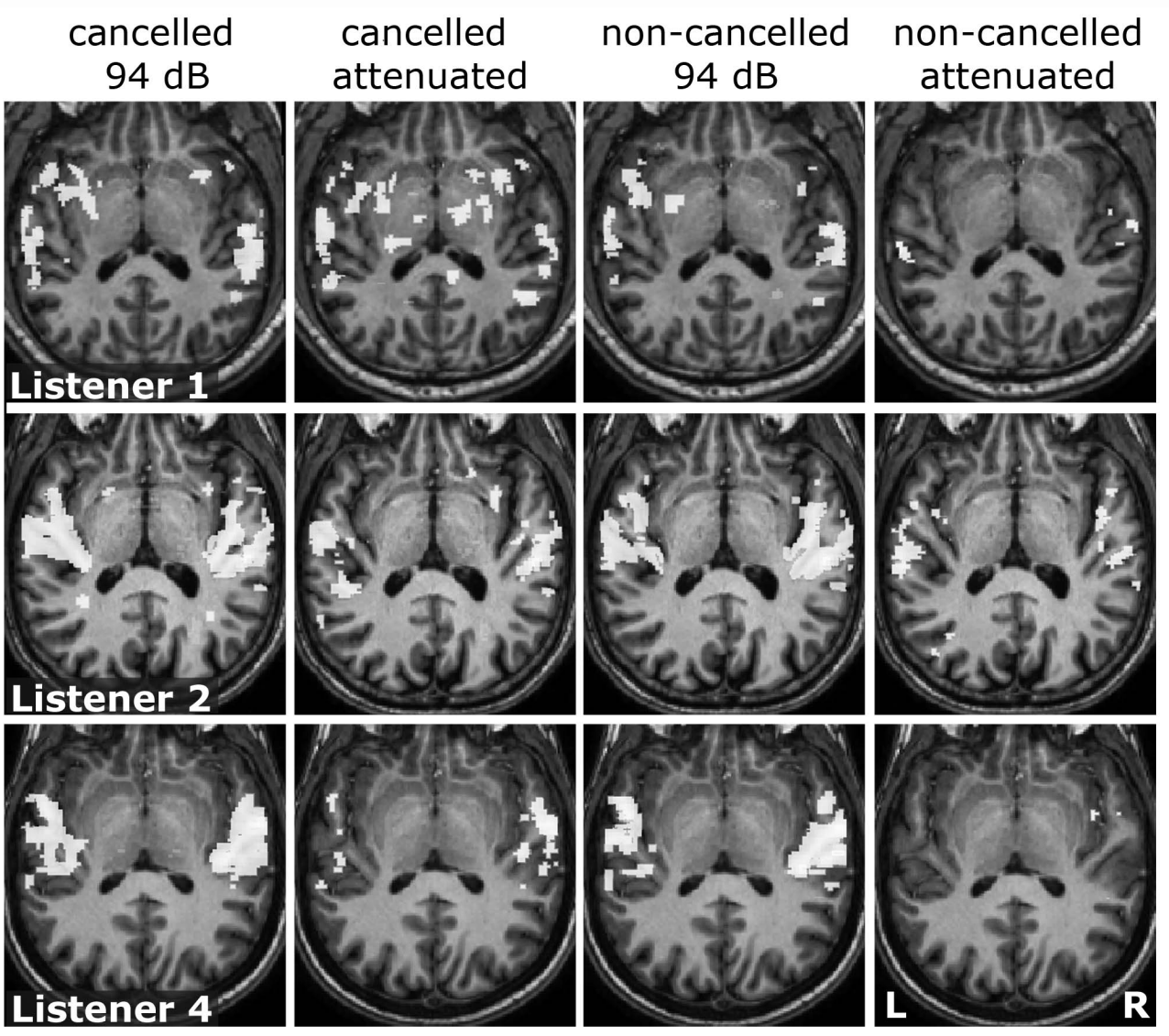
Figure 6. The extent of sound-evoked auditory activity (voxel-level threshold, $p<0.001$) within two auditory cortical regions. The primary auditory region contains a total of 2172 voxels across both hemispheres. The (posterior) non-primary region contains a total of 7982 voxels. Extent of activation is plotted separately for the different sound levels and for the different ANC conditions, with standard error bars.

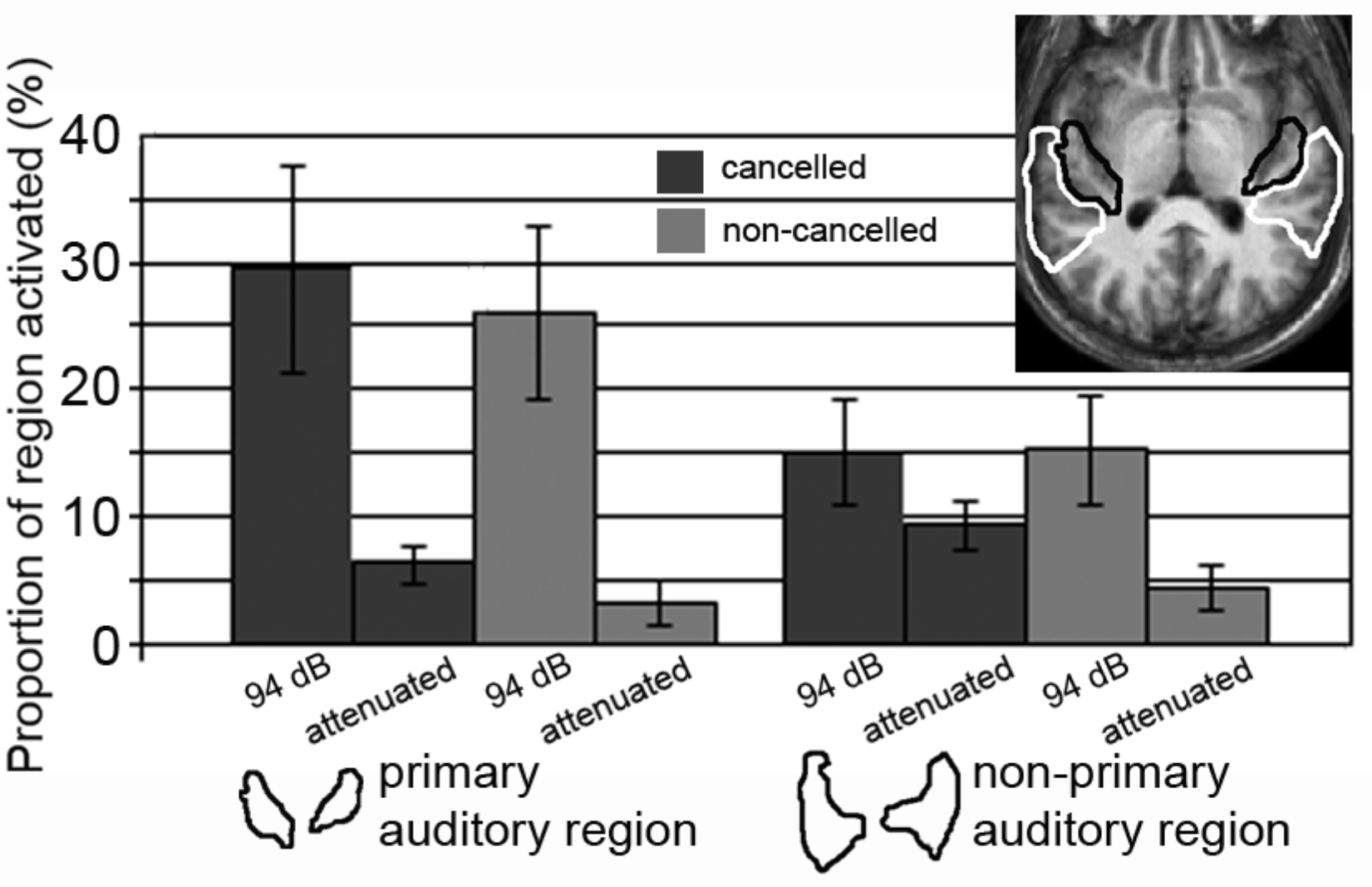

Sound level (dB SPL)
97
85
73
61
49
37
25
non-cancelled
cancelled
1.15 s
55
110
220
440
880
1760
3520
Frequency (Hz)


A
B
RMS level (dB SPL)
99
89
79
69
59
49
0
1
2
3
Time course of the adaptive run (mins)
right ear recording
left ear recording
600-Hz target


non-cancelled masker
cancelled masker 1
cancelled masker 2
Level (dB)
Time (ms)
90
80
70
60
50
40
30
0
300
600
900
1200
1500


Discrimination performance (d')
7
6
5
4
3
2
1
0
cancelled
non-cancelled
94 dB
attenuated
94 dB
attenuated


cancelled
94 dB
cancelled
attenuated
non-cancelled
94 dB
non-cancelled
attenuated
Listener 1
Listener 2
Listener 4
L
R


Proportion of region activated (%)
40
30
20
10
0
cancelled
non-cancelled
94 dB
attenuated
94 dB
attenuated
94 dB
attenuated
94 dB
attenuated
primary
auditory region
non-primary
auditory region