Heschl’s gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes

Marc Schönwiesner1,2*
Nikolai Novitski1,2
Satu Pakarinen1,2
Synnöve Carlson3,2
Mari Tervaniemi1,2
Risto Näätänen1,2
1Cognitive Brain Research Unit, Department of Psychology, University of Helsinki, Helsinki, Finland
2Helsinki Brain Research Centre, University of Helsinki, Helsinki, Finland
3Neuroscience Unit, Institute of Biomedicine / Physiology, University of Helsinki, Helsinki, Finland and Medical School, University of Tampere, Tampere, Finland

Running head: Cerebral pathway for auditory change detection

*Corresponding Author:
Marc Schönwiesner, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, 3801 rue University, H3A 2B4 Montréal, Québec, Canada
marc.schoenwiesner@mcgill.ca

Acknowledgments
We thank Elvira Brattico for helpful discussions throughout the study, Kimmo Alho for comments on the manuscript, and Antti Tarkiainen and Marita Kattelus for technical assistance with the scanner.

Grants
This work was supported by grants #211486, #211487, #211488, and #213933 from the Academy of Finland (National Centre of Excellence Program).

Abstract
A part of the auditory system automatically detects changes in the acoustic environment. This pre-attentional process has been studied extensively, yet its cerebral origins have not been determined with sufficient accuracy to allow comparison to established anatomical and functional parcellations. Here we used event-related functional magnetic resonance imaging and electroencephalography in a parametric experimental design to determine the cortical areas in individual brains that participate in the detection of acoustic changes. Our results suggest that automatic change processing consists of at least three stages: initial detection in the primary auditory cortex, detailed analysis in the posterior superior temporal gyrus and planum temporale, and judgment of sufficient novelty for the allocation of attentional resources in the mid-ventrolateral prefrontal cortex.

Keywords: mismatch negativity, fMRI, event-related potentials, sound duration
Introduction
Detecting changes in the environment is essential for the survival of many organisms. Brain mechanisms of acoustic change detection have been extensively studied in humans using electroencephalography (EEG). The prime experimental model of auditory change detection is the presentation of infrequent deviant events in a stream of repeating standard events. The deviant sounds evoke a frontal negative deflection in the auditory event-related potential, the mismatch negativity (MMN, Näätänen et al. 1978). The MMN can be recorded in response to any discriminable change in the stimulus stream, and the response amplitude correlates with the magnitude of the acoustic change. The MMN is important in two respects: first as a means to study the mechanisms of change detection and how these relate to other cognitive processes like attention and memory, and second as a widely used tool in diverse areas of research, including language acquisition, sound localization, and psychiatric and developmental disorders (Näätänen 1995, 2003). The MMN is often interpreted to imply the existence of a sensory-memory trace in which the features of the frequently occurring standard stimuli are represented. Much research has been dedicated to the translation of this psychological model into neurobiological mechanisms. The localization of the cerebral origin of the mismatch negativity potential was a major aim in several functional magnetic resonance imaging (fMRI), magnetoencephalographic, and high-density EEG studies. Yet, the regional specificity of the results has remained relatively low. Two contributions to the change response, one from the temporal lobes and one from the right frontal lobe, were suggested on the basis of the current density distribution of evoked potentials (Giard et al. 1990) and reductions of the MMN amplitude in patients with lesions in the frontal and temporal lobes (Alain et al. 1998; Alho et al. 1994). Since then, a number of neuroimaging studies have tried to locate the generators of these components (Doeller et al. 2003; Liebenthal et al. 2003; Marco-Pallares et al. 2005; Mathiak et al. 2002; Molholm et al. 2005; Muller et al. 2002; Opitz et al. 1999; Rinne et al. 2005; Schall et al. 2003; Tervaniemi et al. 2006). Results vary substantially across these experiments. Nevertheless, all of these studies report activation in the region of the superior temporal gyrus, sometimes including Heschl’s gyrus, and several found activation of the inferior frontal gyrus.

A more precise localization that allows reliable comparison to known anatomical areas or functional parcellations of the superior temporal and inferior frontal gyri would allow significant progress in the understanding of pre-attentive change detection. For instance, the part of the inferior frontal gyrus in which most of the reported MMN-related activations fall can be subdivided into 3 areas (Brodmann areas 44 and 45, and the deep frontal operculum) with different connection patterns in individual subjects using diffusion tensor imaging (Anwander et al. 2006). A second example is the question whether the change-detection mechanism is co-localized with primary feature processing (i.e. does the detection of small pitch changes happen in areas that extract pitch?). This is a tacit, but unproven, assumption in studies that use the MMN as a tool to locate the processing of acoustic features (Pulvermüller et al. 2006; Shestakova et al. 2004; Tervaniemi et al. 2006).

Several factors have considerably hindered pinpointing the location of the change-related effects with higher regional specificity using fMRI: a) The response to a subtle change in a stream of acoustic stimuli is much smaller in magnitude than the response to an isolated sound; b) the dynamic range of the response is further decreased by the MRI scanner.
noise if images are acquired continuously; c) in block-design experiments, the number of
standard stimuli between deviants is necessarily low, decreasing the amplitude of
responses to the deviants (Haenschel et al. 2005); d) responses to sounds deviating in
frequency (the most commonly used deviant type) are confounded, because infrequent
stimulation with a different frequency will activate ‘fresh’ neural populations in
tonotopically organized cortical areas – this is not the case for other sound features that
are not represented topologically, like violations of temporal order (‘complex MMN’, see
Paavilainen et al. 2001), and perhaps sound duration (but see Pantev et al. 1989); e) studies
using duration deviants to avoid adaptation-related confounds have not accounted
for the decreased stimulus energy of deviants shorter than the standard, further
diminishing the amplitude of the response. Some of the previous studies address one or
more of these points, but none addresses all.

Here we surmount those problems with a parametric event-related fMRI and EEG
experiment, using sparse imaging to eliminate the effects of scanner noise. This
procedure permits localization of the responses in individual brains, as well as individual
comparison of EEG and fMRI results. We measure responses to several deviant
magnitudes to separate different parts of the change detection mechanism.

Material and Methods

Participants

13 volunteers (between 20-30 years of age, 4 male, 10 right-handed) took part in the
current experiment after having given written informed consent. The participants had no
history of audiological or neurological disease. The experimental procedures conformed
with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and
were approved by the Ethics Committee for Ophthalmology, Otolaryngology, Neurology
and Neurosurgery of the University Hospital and by the Ethics Committees of the
Department of Psychology of the University of Helsinki.

Stimuli

All sounds were click trains with a rate of 500 Hz, low pass filtered at 8000 Hz. At this
rate, the clicks are not perceived individually, but as a complex tone with a pitch of 500
Hz, including all harmonics of this frequency up to the low pass filter cutoff. The number
of clicks was varied to generate stimuli of different durations: 100 ms, 74 ms, 52 ms, and
30 ms. The 100 ms click train is in the following referred to as ‘standard’, while the
shorter click trains are referred to as ‘deviants’, specifically as small (74 ms), medium (52
ms), and large deviant (30 ms), according to the acoustical difference from the standard
sound. The sound durations were chosen to elicit an approximately linear increase in the
magnitude of the deviance response (Näätänen et al. 2004). The stimuli were equalized
for root-mean-square energy, so that the energy contour of a sequence of stimuli was
constant over time. Changes in the sound duration were thus the only salient feature in
the stimulus stream to elicit responses. No attempt was made to control for effects of the
physical differences between standard and deviant sounds on brain activation, like the
shorter delay of the neural offset response to deviants of slightly shorter duration than
standard sounds. Such effects can be controlled by reversing the role of standards and
deviants between blocks (Kujala et al. 2006). Note however, that due to its low temporal
resolution, fMRI is relatively insensitive to differences in the temporal layout of the response, such as caused by the slightly different delay of the sound offset. It is therefore unlikely that this particular confound of sound duration changes is a large factor in the observed responses.

During the EEG session, scanner noise recorded from the sequence used in the fMRI experiment was played back to the participants, simulating the acoustical environment in the MRI scanner.

**Procedure**

All participants took part in an EEG recording and in a subsequent fMRI session. During the fMRI session, participants wore pneumatic headphones (which provided sufficient playback quality for the relatively simple stimuli), and looked at a screen though a mirror attached to the head coil. Auditory stimulus presentation was organized in 9 s trials. Each trial started with the 1.2 s sound of the fMRI image acquisition, played back via headphones (in the EEG session) or produced by the scanner (in the fMRI session). Starting 50 ms past trial onset and continuing during the whole duration of the trial 27 100 ms click trains were presented repetitively with a stimulus onset asynchrony of 333 ms. Either 2, 3, 4, 5, or 7 s before trial offset, one of the standard stimuli was replaced by a deviant sound. While irrelevant in the EEG session, this timing allowed estimation of the hemodynamic response to the deviants in the fMRI session (Fig. 1).

We weighted the sampling of the hemodynamic response function according to their potential contribution in locating the responses, i.e. we acquired most repetitions from time points close to the expected peak of the hemodynamic response, thus trading some of the response function estimation power for response detection power. For each deviant, we acquired 15, 20, 20, 20, and 10 repetitions for time points 2, 3, 4, 5, and 7 s, respectively (85 repetitions in total). Additionally, 25 trials containing only standard sounds served as a baseline, making up a total of 16 experimental conditions (3 deviant types × 5 possible onset times within the trial + baseline). Altogether 280 trials (85 trials per deviant × 3 deviant types + 25 baseline trials) were presented in pseudo-random order with equalized transition probabilities. Total experimental time in the fMRI session was 42 min, which was split in 4 runs.

To control attention and direct it away from the acoustic stimuli, and to reduce eye movements, participants were asked to fixate a cross at the center of the screen and perform a visual control task. The task was to press a button with the left or right index finger upon each occurrence of a capital letter in a sequence of random digits that were shown at the center of the screen. The digits were presented for 80 ms at irregular intervals with an average of 4 digits per trial. The target occurred on average once every two trials. After the fMRI session, participants were asked to rate their level of alertness during scanning, the subjective sound level, and the difficulty of the task in comparison to the EEG session.

During the EEG session participants were seated in a comfortable chair in a sound attenuated room. The presentation of the experiment and the task were the same as in the fMRI session. Because the EEG analysis required a higher number of repetitions per deviant, the experiment was run twice during the EEG session.
Electroencephalographic recording and analysis

An electroencephalogram (EEG) was recorded with 128 active sintered Ag-AgCl electrodes (BioSemi, Amsterdam, The Netherlands), positioned radially equidistant from the vertex across the scalp (BioSemi ABC layout). Additional electrodes were placed at the left and right mastoid, at the outer canthi of each eye, at the right eye supra- und infraorbitally, and on the nose tip. The setup does not use a conventional recording reference, but instead actively clamps the average potential of the subject by a feedback loop between two dedicated electrodes to the analog-to-digital conversion reference voltage. The data were recorded direct-current-coupled and digitized with 512 Hz sampling rate. Low pass filtering to avoid aliasing was performed by the decimation filter of the analog-to-digital converter (5th order sinc response, -3 dB point at 102 Hz). The resulting data files were transformed into the Neuroscan continuous data format (PolyRex software [http://psychophysiology.cpmc.columbia.edu/Software/PolyRex], constant gain across all datasets). Signals from the scalp electrodes were re-referenced to the nose tip potential. Signals from the face electrodes were used to compute the horizontal and vertical bipolar electro-oculogram. The data were filtered with a digital band pass filter between 1 and 15 Hz with slopes of 24 dB/octave. Data epochs from 100 milliseconds before to 350 ms post stimulus onset with samples exceeding ±75 µV were rejected from the subsequent analysis. The data were visually inspected for residual artifacts.

Responses to standards (excluding standards directly following deviants) and each of the deviants were averaged separately. The responses to standards were subtracted from those to deviants. In such difference waveforms, the MMN is a negative-going potential at Fz and a positive-going potential at the right and left mastoid in the range of 150 to 250 ms after deviant onset. To assess statistical significance of MMN responses the distribution across participants of peak amplitudes in the baseline interval was compared to the distribution of peak amplitudes across participants in an interval of equal length around the latency of the grand average MMN. This method was chosen instead of the usual comparison of individual peak amplitudes in the MMN latency range to zero, which is slightly biased, because the expected value of amplitude maxima in a certain time range in the absence of a signal is not zero. If this value reflects noise in the signal and the experimental conditions are presented in a balanced pseudo-random sequence, the distribution of baseline amplitude maxima across conditions should be nearly identical. A significant difference in the baseline power between deviant conditions might bias the results of a test for significance of the MMN responses. The peak amplitudes in the root mean square across all channels of the EEG data in both intervals for all participants and deviant conditions were entered into one-tailed paired t-tests. We also tested for equal baseline peak amplitude using a two-way ANOVA with factors ‘interval’ and ‘deviance’. If the baseline power is due to noise, a significant interaction between interval and deviant conditions is expected (effect of deviant in the MMN but not baseline interval). The peak amplitudes of the MMN responses at Fz were additionally subjected to a one-way ANOVA with factor ‘deviance’ and at right and left mastoids with two-way ANOVA with factors ‘deviance’ and ‘hemisphere’. Greenhouse-Geisser correction was applied when necessary.

To obtain a data-driven estimate of the number of components in the MMN response, we performed a spatial principal component analysis of the individual ERPs, and then visualized the probable cerebral origin of the first two principal components as the...
average of the electrode locations weighted by the contribution of each electrode to the PCA component. This estimate takes into account that PCA components are often not dipolar. These location estimates were used solely to seed a subsequent dipole model, and all further analyses and conclusions are based on the dipole analysis. The sources of the responses to the 3 different deviants were analyzed with 3 regional sources, and a four-shell ellipsoidal volume conductor as head model using version 5.1 of the Brain Electrical Source Analysis software (BESA, Gräfelfing, Germany). The locations of two of the regional sources were constrained to be symmetric about the mid-sagittal plane, and fitting was performed within a 50 ms time window centered on the peak of the response. Within the fit window, the residual variance of the model amounted to only 1.4% for the small, 0.7% for medium and 1.1% for large deviant condition. The model was then used as a spatial filter to derive the activation time-course of each regional source (source waveform) for the three deviant conditions in the grand-average and in each individual separately. The orientations of the regional sources were adjusted in each individual so that one of the components captured the maximum of the global field power in the fit window. The peak amplitudes of these components across individuals were analyzed in a two-way ANOVA with factors ‘deviance’ and ‘location’. Greenhouse-Geisser correction was applied when necessary. Post hoc analyses were performed with Tukeys test for honestly significant differences.

**Functional magnetic resonance imaging and analysis**

Blood-oxygen level dependent contrast images were acquired at 3 Tesla (General Electric, SIGNA EXCITE) using gradient echo planar imaging (TR/TE 9000 ms/32 ms) with a head quadrature receiver/transmitter coil. The functional images consisted of 19 ascending slices with an in-plane resolution of 3×3 mm (matrix size 64²), a slice thickness of 3 mm and an inter-slice gap of 1 mm. The 7th slice followed the line connecting the anterior and posterior commissures. The slices were acquired in direct temporal succession in the first 1200 ms of the TR, followed by 7800 ms of stimulus presentation without acquisition noise. This clustering of the slice acquisition at the beginning of a long TR (“sparse imaging”) reduces the effect of scanner noise on the recorded response to the stimuli (Edmister et al., 1999; Hall et al., 1999). A high-resolution structural image was acquired from each subject using a T1-weighted spoiled grass gradient recalled 3D sequence with a resolution of 1 mm³ (matrix size 256×256×150).

Data were corrected for head motion (one participant’s data were excluded from further analysis due to excessive (>2 mm) head translation), spatially smoothed with a 6 mm (individual participant analysis) or 10 mm (group analysis) full-width-at-half-maximum Gaussian kernel, and transformed into the stereotaxic space of the international consortium for brain mapping 152 atlas (‘MNI space’) using the MINC 1.4 software. Statistical analysis was done with Matlab (The MathWorks, Natick, MS) and the FMRISTAT toolbox (Worsley et al. 2002). All deviant conditions were contrasted with the baseline condition and regions of interest (ROI) were defined as 9-voxel neighborhood of local maxima in the resulting statistical parameter map. Hemodynamic response functions in those ROIs were estimated by contrasting responses to each of the 3 deviants at each of the 5 time points separately with the baseline. Because of the high number of modeled conditions, this is the statistically least powerful contrast and only
group level effects are reported. The average hemodynamic response function was used to model the responses to all 3 deviants in a single contrast against the baseline in individual participants. This is the statistically most powerful contrast, and resulting statistical parameter maps were used to identify activated areas in individual brains. The individual results were combined in a random effects analysis to allow inferences on population level (Worsley et al. 2002) and identify activated areas in the group. Signal changes in response to the 3 deviants were extracted from the ROIs to check correlation with EEG results.

Results

Significant mismatch negativity (MMN, Fig. 2) responses to the deviant sounds were observed in EEG recordings (comparison of RMS maxima in baseline and MMN intervals: \( p<0.0001; t_{12}>7 \) for large and medium deviants; \( p<0.0087; t_{12}=2.7 \) for small deviant; ANOVA of RMS maxima in baseline and MMN intervals: effect of interval \( p<0.0001, F_{1,11}=43.9 \), effect of deviant \( p=0.001, F_{2,10}=17.4 \), interaction interval*deviant \( p<0.0001, F_{2,10}=22.6 \), no significant differences in baseline RMS maxima across deviant conditions). MMN amplitudes at Fz and at the mastoids increased with deviance magnitude (ANOVA of voltage maxima: \( p<0.001; F_{2,24}=15.8 \), mastoids: \( F_{2,26}=30.6 \)). There was no significant effect of the factor ‘hemisphere’. The latency of the MMN decreased with deviance magnitude (\( p<0.05; F_{2,24}=4.3 \), mastoids: \( F_{2,26}=4.3 \)).

No differences were observed between the reaction times and detection rates for the visual target detection task between the fMRI and EEG sessions. Most participants reported a similar subjective difficulty of the task in both sessions, but 3 subjects reported that the novel environment of the MR scanner made it initially more difficult to concentrate.

We performed a random effects analysis of the groups’ fMRI data to localize cerebral origins of these responses. Clusters of significant (\( p<0.05 \) corrected) activations were found bilaterally in the temporal lobes and in the right inferior frontal lobe (Fig. 3).

The anatomical loci of the activation maxima were (see Tab.1 for coordinates): medial and lateral Heschl’s gyri (HG), antero-lateral and medial portions of the left planum temporale (PT), portions of the superior temporal gyrus (STG) and sulcus (STS) inferior and posterior to HG, and the mid-ventrolateral prefrontal cortex, bracketed by the horizontal and ascending rami of the inferior frontal sulcus (Brodmann area 45).

Hemodynamic responses in those areas peaked about 3-4 s post deviant onset and returned to baseline about 7 s after deviant onset (Fig. 4, upper panel).

A hemodynamic response function was fitted to the average of all responses and used to model the response magnitude for each deviant across the 5 time points. In activated
areas of the superior temporal lobe, but not in the prefrontal cortex, the response magnitude increased with the deviance magnitude (Fig. 4, lower panel).

Results of individual participants
To check the consistency of the group results across individuals, statistical parameter maps were computed for each participant. Activated areas were compared to individual structural images, and effect sizes were extracted from local maxima in the maps. The constant stream of standard stimuli reduces the dynamic range of responses to deviant stimuli compared to typical sound versus silence contrasts. Responses were nevertheless clear in all participants, albeit with a high degree of inter-individual variability (Fig. 5).

Note that while the locations of the suprathreshold local maxima were variable, the regions that show responses just below threshold were similar in all participants (leading to significant activations in the random effects analysis). The majority of the activation foci in the superior parts of the temporal lobes were in HG, STG, STS, and PT. The small foci in the frontal lobes were consistently located in parts of the mid-ventrolateral prefrontal cortex across subjects.

To compare the responses from temporal and frontal sites between fMRI and EEG sessions, we separated the individual ERPs into temporal and frontal components using source analysis. To check whether this separation is in fact possible in a data-driven manner without knowledge of the fMRI activation sites, we performed a spatial principal component analysis of the individual ERPs. In all individuals, the first component was localized between the superior parts of the left and right temporal lobes, consistent with a superposition of bilateral responses from the auditory cortices. The second component was located in the inferior part of the frontal lobe in the majority of individuals, with a slight average lateralization to the right. The locations of these two components indeed suggested temporal and frontal contributions to the overall response, consistent with the hemodynamic response pattern.

In order to obtain a reliable estimate of the response of the temporal and frontal contributions to the different deviant sounds, we analyzed the sources of the evoked potentials with equivalent current dipole modeling, using the location estimates of the principal components as seeds. The group average responses were modeled with 3 regional sources, 2 of them with a symmetry constraint to account for bilateral responses from the auditory cortices. The resulting sources were located bilaterally in the vicinity of HG and in the right frontal lobe. The spatial resolution of this model is relatively low, but the general locations agree with the fMRI data. The model was then used as a spatial filter to derive the activation time-course of each regional source (source waveform) for the 3 deviant conditions in the grand average (Fig. 6) and the peak amplitudes in each individual separately (Fig. 7, left panel).
In both EEG and fMRI data, the response magnitudes from sites in the temporal lobes increased with deviance magnitude, whereas the responses from frontal sites were not modulated by the deviance magnitude across participants ($F_{2,11}>30$, $p<0.0001$ for effects of response origin, deviance magnitude, and their interaction; post-hoc tests for deviance dependence of temporal sites $p<0.001$, and frontal sites $p=0.89$) (Fig. 7). Activity of the frontal sites showed in some individuals a trend towards increasing or decreasing responses with deviance magnitude. Of the responsive sites in the temporal lobes, in 9 out of 12 participants the posterior STG and neighboring lateral PT showed the clearest modulation by deviance magnitude.

Discussion

Using an experimental design that overcomes previous methodical difficulties, we were able to localize the cerebral sites involved in the detection of duration changes in a constant stream of acoustic stimuli with greater precision than previously achieved. We characterize the degree of inter-individual variability and demonstrate, at the level of individual brains, different response patterns of temporal and frontal sites in the high-density EEG and fMRI data.

Responses in the temporal lobes were found in lateral and medial portions of HG, in the medial and lateral PT bordering HG, and along STG and STS, mostly posterior to HG. In earlier brain imaging studies on preattentive auditory deviance detection, activation of the STG has been the most consistent finding (Doeller et al. 2003; Liebenthal et al. 2003; Mathiak et al. 2002; Molholm et al. 2005; Muller et al. 2002; Opitz et al. 1999; Opitz et al. 2002; Rinne et al. 2005; Schall et al. 2003; Tervaniemi et al. 2000; Tervaniemi et al. 2006). The region responsive to deviants in our study encroached the PT, mostly adjacent to STG, STS, and medial and lateral parts of HG. The posterior STG and adjacent parts of the planum temporale showed an increase in activity with increasing deviance magnitude. Several anatomical areas have recently been delimited on the STG using observer-independent measures of differences in cyto- and receptoarchitecture (Morosan et al. 2005; Schleicher et al. 2005). According to this schema, area Te3 covers the posterior two thirds of the outer convexity of the STG (a posterior portion of Brodmann area 22). Its location fits well with the activated regions of STG that showed the highest dependence on deviance magnitude. Homologue regions in the primate auditory cortex belong to the parabelt, a tertiary region in the processing hierarchy of the auditory cortex (Kaas and Hackett 2000) that receives indirect connections from the primary auditory cortex via adjacent secondary (‘belt’) areas (Kaas and Hackett 1998). In humans, the belt region would occupy the lateral HG, and parts of the PT and planum polare adjacent to HG (Galaburda and Sanides 1980). We indeed found responses to the deviant sounds in the lateral HG and the planum temporale.

Medial portions of HG were found active in 6 individuals (9 if responses slightly below statistical significance are included) out of 12 participants. This suggests a contribution of the primary auditory cortex to deviance detection, which is in agreement with the demonstration of stimulus-specific adaptation of responses in the primary auditory cortex – a candidate neural correlate for some of the change responses observed in humans (Ulanovsky et al. 2003). Recordings of duration-MMN responses from depth electrodes in human HG also implicate the primary auditory cortex in change detection (our unpublished observations). Opitz and colleagues (2005) demonstrated that, during
detection of frequency deviants, secondary areas on lateral HG seem to mediate a memory-trace based mismatch response, whereas activity on medial HG is related to a sensory mechanism of change detection, the stimulation of non-refractory portions of tonotopic auditory cortex. The authors conclude that both of these mechanisms contribute to the MMN evoked by frequency deviants. It is unclear whether a similar sensory mechanism can account for the activation of primary auditory cortex during the detection of duration deviants. In mammals, duration-selective neurons have been found in the mouse inferior colliculus (Brand et al. 2000) and cat auditory cortex (He et al. 1997), but there is no indication of a large-scale topographic representation of sound duration in the human primary auditory cortex.

Based on these findings, we suggest that changes in the acoustic environment are initially detected at or below the level of the primary auditory cortex. Because the responses from the posterior STG and lateral PT follow the deviance magnitude most closely, these structures might extract the details of the acoustic change. The STS may be involved in a secondary process that is only loosely dependent on deviance magnitude and relies on input from the STG. There is indeed evidence that activity in the STS is more dependent on involuntary shifts of attention towards the deviant sound than on passive detection of the deviant (Sabri et al. 2006). Interestingly, the hemispheric distribution of the responses to duration changes might depend on the nature of the stimulus. In the present study, we found no significant systematic differences in the responses from the left and right superior temporal plane, whereas Tervaniemi and colleagues (2006) demonstrated that complex speech- and music-like stimuli might evoke stronger responses in the left and right STG/STS, respectively.

In the majority of the participants, the responses in the IFG were found between the ascending and horizontal ramus. This region of the mid-ventrolateral prefrontal cortex corresponds to Brodmann area 45 (Amunts et al. 1999), and is connected with the STG and STS via the arcuate and superior longitudinal fascicle (Geschwind 1970; Petrides and Pandya 2002). There was a latency difference of about 50 ms between the peak in activity of the frontal and temporal sources in our equivalent current dipole model of the EEG responses. Tse and colleagues (2006) found a similar latency difference of about 60 ms with optical imaging in humans between activity in the superior temporal and inferior frontal lobe. These differences in the latency of temporal and frontal responses suggest that change-related activity in the IFG relies on afferent projections from the perisylvian region of the temporal lobes. Moreover, the latency difference is almost an order of magnitude higher than the passive conduction time between the two sites, suggesting that the activity observed in the mid-ventrolateral prefrontal cortex indicates cerebral processing rather than passive conduction. This processing is thought to relate to a possible switch of attention to the deviant sound (Giard et al. 1990). The P3a (Squires et al. 1975) is an event-related potential, probably generated in the prefrontal cortex (Knight 1984), and thought to indicate the allocation of attentional resources to novel events (Daffner et al. 2000; Escera et al. 1998). The latency of the frontal source (just between MMN and P3a), and its independence of deviance magnitude (whereas MMN and P3a increase in amplitude with deviance magnitude (Näätänen et al. 2004; Schröger and Wolff 1998)) suggest that activity of the frontal source might indicate an intermediate step, perhaps a decision whether the stimulus is sufficiently novel to require attentional resources. Our deviant sounds are presented 85 times each and probably cease to be novel
to the participants after the first few presentations. Because novelty is not an attribute of the stimulus, but an evaluation that arises in the brain of the listener, it must have a neural correlate. While the detection of a deviating sound can be based on sensory memory (which decays within a few seconds, Mäntysalo and Näätänen 1987; Sams et al. 1993), the decision whether this sound is in fact novel (was not heard previously during the experiment) would have to be based on a process with a longer lasting memory span than the MMN system. An involvement of the right mid-ventrolateral prefrontal cortex (Brodmann areas 45 and 47/12) in such memory-based decisions has been demonstrated in humans and macaque monkeys (Petrides 2005). Petrides and colleagues (2002) found activation of the mid-ventrolateral prefrontal cortex when human participants had to find a novel stimulus in pairs of familiar and novel visual stimuli. Activation in the perisylvian areas of the temporal lobe related to the detection of a deviant stimulus may trigger activity in the mid-ventrolateral prefrontal cortex to quickly determine whether the deviant stimulus requires additional attentional resources (‘call for attention’ (Öhman 1979)). However, at least one study found frontal activity preceding activation in the temporal lobes by about 20 ms (Yago et al. 2001). According to these authors, the early frontal activity might reflect either a genuine MMN component, or a frontal subcomponent of the N1 involved in the processing of the frequency deviants presented, or an artifact of the procedure. Nevertheless, because activity in temporal and frontal regions largely overlaps in time, it is possible that one response is not merely triggered by the other, but that information is flowing back and forth between these areas.

In summary, our results suggest that at least three regions of the cerebral cortex are involved in the automatic processing of acoustic changes: the primary auditory cortex, the posterior superior temporal gyrus and planum temporale, and the mid-ventrolateral prefrontal cortex. Analysis of the timing of activity in the EEG data and comparison to previous results support a hierarchical model in which these three regions are involved in the initial detection of an acoustical change, a detailed analysis of the change, and judgment of sufficient novelty for the allocation of attentional resources, respectively.
References


Figure 1: Experimental design in the fMRI session. The image acquisition (grey bars) left 7.8 s of silence for the stimulus presentation (row of black lines) and decay of the response to the scanner noise (dashed line). Deviant stimuli (black arrows) were presented at different time points in relation to the image acquisition. This allowed the sampling of different time points (white arrows) of the hemodynamic response (black curve) at 2, 3, 4, 5, and 7 s (from right to left) after deviant onset.
Figure 2: Grand-average mismatch potentials in response to the large, medium, and small deviants at different locations across the scalp (nomenclature of locations according to the extended 10-20 system, Jasper 1958; Sharbrough et al. 1991). The amplitude difference between responses to standard and deviants is plotted over time relative to the stimulus onset.
Figure 3: Activations in response to duration changes. Statistical parameter maps showing significant responses (p<0.05 corrected) shown on sections through the mean structural image of the group (A-E, radiological orientation, inset indicates slice locations) and on an individual gray matter surface (international consortium for brain mapping single subject anatomical template; right [F], and left hemisphere [G]). HG Heschl’s gyrus, STG superior temporal gyrus, PT planum temporale, STS superior temporal sulcus, PFC mid-ventrolateral prefrontal cortex.
Figure 4: Hemodynamic responses from different cerebral locations to the small, medium, and large deviants (arbitrary units). HG Heschl's gyrus, PT planum temporale, STG superior temporal gyrus, STS superior temporal sulcus, IFG ar/hr ascending and horizontal rami of inferior frontal gyrus
Figure 5: Individual activation patterns (radiological orientation, position and extent of slice portions indicated at bottom). Responses in the superior temporal lobes showed considerable variation across the 12 participants. Responses in the frontal lobes (arrows) clustered consistently around the horizontal and ascending rami of the inferior frontal gyrus.
Figure 6: Activation time courses obtained from the equivalent current dipole model of the grand-average EEG responses to the 3 deviants. The responses from the bilateral temporal sources strongly depend on deviance magnitude. They are followed about 50 ms later by the response from the right frontal source that is independent of deviance magnitude.
Figure 7: Comparison of individual temporal-lobe (black bars) and frontal-lobe (white bars) responses between EEG and fMRI sessions. For clarity, only the fMRI activity of the local maximum that showed closest correlation with the EEG responses is shown. r/l left/right side structure, HG Heschl's gyrus, PT planum temporale, STG superior temporal gyrus, STS superior temporal sulcus, IFG inferior frontal gyrus.

<table>
<thead>
<tr>
<th>#</th>
<th>EEG</th>
<th>FMRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>ISTG</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#2</td>
<td>rSTG</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#3</td>
<td>ISTG</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>rPT</td>
<td>rIFGsup/IFGant</td>
</tr>
<tr>
<td>#5</td>
<td>IPI</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#6</td>
<td>rPT</td>
<td>rIFG</td>
</tr>
<tr>
<td>#7</td>
<td>rPT</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#8</td>
<td>rHG</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#9</td>
<td>rSTG/rSTS</td>
<td>rIFGantsup</td>
</tr>
<tr>
<td>#10</td>
<td>rSTS</td>
<td>rIFGsup/ant</td>
</tr>
<tr>
<td>#11</td>
<td>rHG</td>
<td>rIFGant</td>
</tr>
<tr>
<td>#12</td>
<td>ISTG</td>
<td>rIFGant</td>
</tr>
</tbody>
</table>

60 nAm 1.4 %

<table>
<thead>
<tr>
<th>temporal</th>
<th>frontal</th>
</tr>
</thead>
</table>

- ISTG: Intra-hemispheric single-trial generalized temporal activity
- rSTG: Right superior temporal gyrus
- rPT: Right planum temporale
- rHG: Right Heschl's gyrus
- rSTG/rSTS: Right superior temporal gyrus and planum temporale
- rSTG/rSTG: Right superior temporal gyrus
- rIFG: Right inferior frontal gyrus
- rIFGsup/IFG: Right superior/inferior frontal gyrus
- 60 nAm: Amplitude level
- 1.4 %: Percentage of activity
<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>coordinates (x,y,z)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>HG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>medial</td>
<td>L</td>
<td>-44, -25, 7</td>
<td>6.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>lateral</td>
<td>L</td>
<td>-52, -12, -1</td>
<td>6.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>medial</td>
<td>R</td>
<td>42, -20, 5</td>
<td>6.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>lateral</td>
<td>R</td>
<td>46, -10, 1</td>
<td>5.6</td>
<td>0.003</td>
</tr>
<tr>
<td>PT</td>
<td>L</td>
<td>-50, -35, 16</td>
<td>6.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>STG</td>
<td>L</td>
<td>-63, -28, 10</td>
<td>8.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>65, -36, 6</td>
<td>5.7</td>
<td>0.002</td>
</tr>
<tr>
<td>STS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>middle</td>
<td>L</td>
<td>-64, -16, -2</td>
<td>6.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>anterior</td>
<td>L</td>
<td>-62, -2, -8</td>
<td>6.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>60, -23, -4</td>
<td>5.3</td>
<td>0.001</td>
</tr>
<tr>
<td>IFG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ascending ramus</td>
<td>R</td>
<td>46, 23, 11</td>
<td>4.9</td>
<td>0.041</td>
</tr>
<tr>
<td>horizontal ramus</td>
<td>R</td>
<td>49, 32, 2</td>
<td>5.0</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Table 1: Brain areas activated by duration changes in the acoustic stimulus stream. Coordinates are in MNI space (lookup available at www.bic.mni.mcgill.ca/cgi/icbm_view); P values are corrected for multiple comparisons. HG Heschl’ gyrus, PT planum temporale, STG superior temporal gyrus, STS superior temporal sulcus, IFG inferior frontal gyrus.