Brain activity underlying auditory perceptual learning during short period training: simultaneous fMRI and EEG recording

Ana Cláudia Silva de Souza1*, Hani Camille Yehia2, Masa-aki Sato3 and Daniel Callan3

Abstract

Background: There is an accumulating body of evidence indicating that neuronal functional specificity to basic sensory stimulation is mutable and subject to experience. Although fMRI experiments have investigated changes in brain activity after relative to before perceptual learning, brain activity during perceptual learning has not been explored. This work investigated brain activity related to auditory frequency discrimination learning using a variational Bayesian approach for source localization, during simultaneous EEG and fMRI recording. We investigated whether the practice effects are determined solely by activity in stimulus-driven mechanisms or whether high-level attentional mechanisms, which are linked to the perceptual task, control the learning process.

Results: The results of fMRI analyses revealed significant attention and learning related activity in left and right superior temporal gyrus STG as well as the left inferior frontal gyrus IFG. Current source localization of simultaneously recorded EEG data was estimated using a variational Bayesian method. Analysis of current localized to the left inferior frontal gyrus and the right superior temporal gyrus revealed gamma band activity correlated with behavioral performance.

Conclusions: Rapid improvement in task performance is accompanied by plastic changes in the sensory cortex as well as superior areas gated by selective attention. Together the fMRI and EEG results suggest that gamma band activity in the right STG and left IFG plays an important role during perceptual learning.

Keywords: Neural plasticity, Attention and performance, Perceptual learning, Auditory perception, Simultaneous fMRI and EEG, Time-frequency analysis

Background

The fact that cortical representations in adult animals can be modified by experience has led to extensive research regarding the neurophysiological mechanisms of cortical plasticity [1,2]. It is apparent that the knowledge of how plasticity can be induced would be of great value in developing treatment for individuals with brain damage or to optimize learning strategies in a normal brain. The capacity of reorganization, at least partly, accounts for certain forms of learning. Learning comes in many forms, some of which are explicit memories of objects, sounds, events and some of which are implicit and nondeclarative. One form of implicit memory, perceptual learning, involves improving one’s ability with practice, to discriminate differences in the attributes of simple stimuli.

One of the most interesting aspects of human sensory perception is that it is not restricted to an early critical life period but can be improved with practice even in adulthood [3]. Relatively little is known about how practice influences the performance of human adults on basic discrimination tasks but the understanding of the physiological substrates of learning will help the development of perceptual training schemes. Most of the perceptual learning studies are directed to the visual system. A number of studies have worked on primitive visual features such as hyperacuity and contrast discrimination [4,5], orientation [6-8], direction of motion [9,10] and texture discrimination [11].

Compared with the investigations in the visual system, the examination of perceptual learning in the auditory system is still in maturation. In traditional psychoacoustic...
experiments, training has been used mainly for the purpose of reaching asymptotic performance. More recently in the literature of learning in the auditory system, there has been an increase of the potential application of auditory training in the treatment of communication disorders [12-14], perceptual expertise [15-17], rehabilitation of abnormal perception [18,19] and improvement of cognitive skills [20-22].

One important aspect of perceptual learning involves its relation to the amount of training. According to Demany [23] few weeks of practice and many trials may be necessary to reach an individual’s asymptotic discrimination threshold. However, recent research indicates that substantial perceptual learning may occur in the very first trials, as evidenced by the improvements made early in learning by participants [24-27]. Another feature that influences learning tasks is the daily limits of learning. Wright and Sabin [28] observed that training beyond some amount in a single day does not increase the amount of improvement. Therefore, whilst traditional approaches work with long term training, it is important to incorporate early trials into perceptual learning experiments rather than just ignoring them. Although it is accepted that slow perceptual learning is accompanied by enhanced stimulus representation in sensory cortices [29,30], the neural substrates underlying early and rapid improvements are still not fully understood. Recent studies suggest that increased accuracy during the first hour of training may involve increased perceptual sensitivity [31]. Alain et al. [29] showed that the perception of two vowels presented simultaneously could be improved within 1 hour of practice and that improvement coincided with enhancements in an early evoked response (~130ms) localized in the right auditory cortex and a late evoked response (~340ms) localized in the right anterior superior temporal gyrus as well as the inferior prefrontal cortex. Moreover, these learning-related changes were restricted only to participants who attended to the task. The importance of attention in perceptual learning has been reported in many studies as well [21,32-35]. During auditory frequency discrimination, attention seems to play an important role in the process underlying complex auditory tasks, such as comprehension and understanding [36-38]. However, as Jagadeesh [1] discussed in his review it is also possible that plasticity happens in the absence of attention. In this case learning may rely on the inherent salience of the stimulus used to induce plasticity. Attention is drawn implicitly by the stimulus, rather than managed consciously by the individual. Some examples of this type of passive perceptual learning are given in [39] and [40].

To our knowledge, cognitive experiments have investigated changes in brain activity after relative to before perceptual learning. However, brain activity during perceptual learning has not been explored. We used electrophysiology EEG and functional magnetic resonance imaging fMRI to examine the brain alterations related to fast perceptual learning. In this study we investigate the extent to which enhanced perceptual discrimination results in greater brain activity in modality specific cortex (auditory) to the perceptual event and to what extent frontal regions participate in prediction and top-down modulation of auditory selective attention that gives rise to auditory perceptual learning. For this purpose we designed a paradigm to test auditory frequency discrimination performance during rapid training in which the level of difficulty was based and controlled by an adaptive staircase method. Applying simultaneous EEG and fMRI recording as well as behavioral data, we are able to investigate the underlying sources of activation related to the course of perceptual learning.

**Methods**

**Subjects**

Simultaneous EEG/fMRI recordings were obtained from 11 subjects (10 males), 22 to 40 years old (mean age 24 years old), with no auditory or visual complaints. Each participant provided informed written consent to participate in the study, which was conducted in accordance with institutional ethical provisions and approved by ATR Human Subject Review Committee in compliance with the Declaration of Helsinki.

**Auditory stimulus**

Each auditory stimulus was composed of five tones (400Hz, 600Hz, 700Hz, 800Hz and 1000Hz) with a total duration of 150ms (10ms of rise and fall times) and loudness level of 90 dB SPL. A deviant stimulus differed from the standard in the frequency of the fourth tone. Frequency deviations varied from 1Hz to 40Hz with steps of 1Hz. A sequence of five stimuli was delivered with random ISI ranging from 450 to 500ms. Each sequence had at most one deviant sound on positions 2, 3, 4 or 5. Stimuli were delivered binaurally through a plastic tube attached to foam earplugs using an MRI/EEG compatible system. The tube introduced a constant delay of 64ms in sound presentation to the ears.

**Visual stimulus**

Visual stimulus followed the same paradigm. The standard stimulus consisted of a white rectangular horizontal bar positioned in the center of the screen (40cm from the eyes viewed through a mirror). The deviant bars were also positioned in the center but rotated clockwise in steps of 0 to 12 degrees. Stimuli were delivered in sequences of five separated by 450 to 500ms. As in the auditory stimulus presentation, in each sequence of five, there was only one deviant bar and it was never in the first position.

**Behavioral test**

Frequency and position discrimination thresholds were measured for each subject in the auditory and visual
conditions, separately, in a sound attenuation booth of 40 dBA. The frequency difference between the deviant tones in each trial was changed in a one-up two-down staircase procedure. A staircase is a procedure in which the order of stimulus presentation is determined by responses given by the listener to the trials that were presented previously. In a frequency detection task it provides a method of estimating the signal level that is required for the subject to obtain a particular proportion of correct responses. Therefore, a one-up two-down staircase targets the 71% correct performance level on the psychometric function [41]. In this method the stimulus level is decreased after two positive responses or increased after one negative response in each trial. A positive response requires correctly detecting a deviant in a sequence of five sounds or five bars (in case of visual stimuli). At the end, threshold estimation was done using the arithmetic mean of reversal values [42]. In the visual test, the ability to determine small variations in clockwise rotation of a rectangular bar from horizontal position was tested. The discrimination level obtained in the behavioral test was used as a starting point for the staircase in the MRI experiment.

3D scanning
After the behavioral test, a 64 channel electrode cap (BrainCap-MR 64 BrainProducts, Munich, Germany) was placed on the subject. A three dimensional (3D) digitizer (FastScan hand-held laser scanner) was used to acquire subject’s head shape and each electrode’s position. Surface volumes were later used for source localization procedures.

Cortical surface model
A polygon cerebral cortex model was constructed using the MRI T1 structural image for each subject. The cortical model assumes a current dipole at each vertex at which the fMRI activity elicited by the stimulus exceeded a threshold. The dipole current directions are assumed perpendicular to the cortical surface [43]. Moreover, subjects’ head shapes obtained from the 3D scanner and the structural images were fit using a least squares method. The head was segmented into three compartments: skin, skull and cerebrospinal fluid. Such segmentation was done in Curry software using the boundary element method.

fMRI experimental design
In the main experiment EEG and fMRI were recorded simultaneously. Stimuli were delivered based on the same staircase procedure used in the behavioral test. A sparse image acquisition technique was applied to prevent contamination of the blood oxygenation level dependent (BOLD) response by the acoustic noise of the scanner and to limit the epochs of contamination of the EEG by the gradient switching during the image acquisition. Functional MRI data were acquired using a Shimadzu Marconi’s Magnex Eclipse 1.5T PD250 scanner. Functional data consisted of T2*-weighted, gradient echo, echo-planar imaging sequence (TE=48ms and flip angle 90°). During each scan, 165 volumes were acquired over 16.5min. The repetition time (TR) was 6 seconds and the scanning time (TA) was two seconds. Stimulus presentation was made during the “silent” four seconds period. Each volume was composed of 20 axially oriented contiguous slices with 4×4×5mm voxel dimensions with 1mm gap between slices. fMRI data from the first two volumes of each session were discarded to avoid the effects of magnetic saturation. At the end of the experiment a T1-weighted structural scan was acquired to align functional data across multiple runs to the subject’s reference volume.

The experiment was composed of two types of task conditions: auditory and visual. Trials of a single condition were grouped together in blocks of 18 sequences of ten stimuli (five auditory and five visual) lasting 120 seconds in total. Auditory and visual stimuli were interleaved in a sequence separated by a pseudo-random interval ranging from 150 to 175ms. Each block started with a visual instruction in the center of the screen 40cm far from the subject’s eyes. Based on what was shown (–Picture of an ear for auditory condition- or -Picture of an eye for visual condition-) the subject had to pay attention to the auditory or visual stimuli. Each instruction lasted four seconds on the screen. Task order was counterbalanced across scanning runs and subjects. Stimuli were delivered during the four seconds of silence when there was no scanning. Before each sequence of stimuli there was a baseline ranging from 650ms to 800ms. After each sequence of 10 stimuli (five visual and five auditory), participants were asked to indicate, by pressing a button (after a green cross appeared on the screen) whether or not a deviant signal was present in the sequence. In this experiment, ‘No’ responses can be either without deviant or with deviant below subject’s perceptual level. A happy face was provided for correct responses, whereas a sad face was presented for incorrect responses. There was a rest condition after each instruction as well as at the end of each block. Figure 1 shows a scheme of the experiment. The recording session consisted of four runs of eight blocks each (four blocks of auditory attention and four blocks of visual attention), resulting in 144 trials acquired per condition per run, with short breaks between them. In this experiment, non-attention to stimulus was attained drawing subject’s attention to the other modality (visual or auditory).

EEG recording
EEG (64-channel) was acquired simultaneously using the Brain Amp MR+fMRI-compatible recorder system in a
continuous mode and the BrainCap-MR 64 electrode cap. Potentials recorded at each site were referenced to the center of the head (Cz). Eye movement activity was monitored with an electrode below the left eye. ECG was also recorded simultaneously. The electrode resistance was kept below 5 kΩ and the data was sampled at 5 kHz per channel.

Functional image analysis
Analysis was carried out using SPM2 (Wellcome Trust Centre for Neuroimaging, UK). This version was chosen because of the compatibility with VBMEG (source localization procedure). Preprocessing was performed on functional and anatomical images using a common procedure: slice timing, movement correction, normalization and smoothing. Subjects’ functional images were coregistered to their own anatomical T1 images. Images were spatially normalized to a standard anatomical space defined by a template T2 image from the MNI (Montreal Neurological Institute), resampling every 3 mm using sinc interpolation. Finally, functional images were smoothed with an 8 mm FWHM (full-width half maximum) Gaussian kernel. Brain activation during experimental conditions was estimated for each subject using event related fMRI, based on the onset of individual events in the general linear model. Statistical parametric maps were generated for each subject for each experimental condition: auditory response in auditory task (stimulus attended); auditory response in visual task (stimulus unattended) and rest period. Significant voxel activation was determined using $t$-statistics with a threshold of $p<0.005$, uncorrected.

To localize brain regions involved in attention demands, activations in the attended and unattended conditions were directly contrasted. In addition, a measure of performance change indicating learning was assessed using the difference between beginning and ending thresholds as a regressor in each session for the auditory-attended condition. It was not possible to investigate the attention related learning effect by doing the analysis over the contrast of the auditory-attended relative to the auditory unattended condition because the auditory unattended condition corresponded to the visual-attended condition in which visual learning was taking place. It becomes somewhat complex to run the modulation of both auditory and visual learning components when learning effects are occurring for both aspects of the contrast of auditory-attended relative to visually-attended (auditory-unattended). Therefore we ran the learning related modulation over the auditory-attended condition only, without subtracting out the visually-attended condition first. To account for performance related variability across subjects, the design matrix was weighted (simple regression analysis) with each subject’s overall gain in a second level analysis.

EEG data preprocessing
In this study the artifact template subtraction proposed by Allen et al. [44] was used to remove the gradients produced by the switching of magnetic gradients. This approach assumes that the shape of gradient artifacts is constant over time and additive to the physiological signal. Subsequently, independent component analysis (ICA) was conducted over the epoched and baseline

---

**Figure 1** Schematic description of the experimental design.
removed data (650ms prior to and 3075ms after stimulus onset) in order to extract ballistocardiogram, ocular and movement artifacts [45,46]. The rejection of components was determined by finding a cross-correlation (Pearson’s r>0.3) between each IC and the electrooculogram (EOG) as well as the electrocardiogram (ECG) channels recorded simultaneously with neuronal data. Rejection was also carried out based on abnormal linear trends (using a window width of 932 points, maximum acceptable slope of 0.5 and coefficient of determination R²> 0.3). As a final criterion, rejection was carried out by inspecting the components topographic scalp map for characteristics of normal artifact such as eye movement, eye blinks and muscle activity.

The variational hierarchical Bayesian method was used to constrain EEG inverse solutions to regions where fMRI indicates large hemodynamic activation [43,47]. For the estimation, EEG data were divided into 600ms windows with 85% overlap. The prior for each time window was given by the fMRI activity corresponding to the stimulus shown during that time window. The hyper-parameters that control the relative amplitude of the prior current variance and the width of the prior distribution were set m₀=100 and γ₀=100. The current variance estimation was done using the time sequence of all trials. Each individual’s fMRI activity of all experimental conditions (auditory task attended and unattended) was used as a source localization constraint. For single trial current estimation, the Bayesian inverse filter was applied to three areas of interest determined by using a current estimation, the Bayesian inverse filter was applied as a source localization constraint. For single trial conditions (auditory task attended and unattended) was applied to 6mm radius spherical regions of interest (ROIs) as well as the electrooculogram (EOG) channels recorded simultaneously with neuronal data. Rejection was also carried out based on abnormal linear trends (using a window width of 932 points, maximum acceptable slope of 0.5 and coefficient of determination R²> 0.3). As a final criterion, rejection was carried out by inspecting the components topographic scalp map for characteristics of normal artifact such as eye movement, eye blinks and muscle activity.

The variational hierarchical Bayesian method was used to constrain EEG inverse solutions to regions where fMRI indicates large hemodynamic activation [43,47]. For the estimation, EEG data were divided into 600ms windows with 85% overlap. The prior for each time window was given by the fMRI activity corresponding to the stimulus shown during that time window. The hyper-parameters that control the relative amplitude of the prior current variance and the width of the prior distribution were set m₀=100 and γ₀=100. The current variance estimation was done using the time sequence of all trials. Each individual’s fMRI activity of all experimental conditions (auditory task attended and unattended) was used as a source localization constraint. For single trial current estimation, the Bayesian inverse filter was applied to three areas of interest determined by using a mask with the learning contrast and extended voxels equal to 50 to clear out areas of no interest.

**Results**

**Behavioral data**

Behavioral data acquired during the experiment shows an exponential, quasi-linear and decreasing tendency in perceptual auditory frequency discrimination thresholds (r=0.99, p=0.0041). Figure 2 shows the grand mean and deviant error of 11 subjects. Although we have used a similar experimental paradigm for the auditory and visual conditions, no behavioral learning effect seems to happen as shown in Figure 3. Given the lack of any behavioral learning effect it is unlikely that the visual stimuli would evoke a visual learning response.

**Functional magnetic resonance imaging**

The brain imaging results of the auditory attended relative to rest contrast show activation in the temporal, frontal and parietal cortices. The auditory unattended (visual attended) relative to rest condition shows activation in parietal, occipital and temporal cortices as summarized in Table 1. Statistical parametric maps for these conditions are given in Figure 4A-B (Auditory: T=2.49, p_FDR<0.05, spatial extent threshold=90 voxels; Visual: T=2.66, p_FDR<0.05, spatial extent threshold=90 voxels; spatial extent is selected based on uncorrected cluster level p<0.05).

With regards to evaluating the attentional load on the task, a direct contrast between auditory attended and auditory unattended (visually attended task) conditions was conducted using the intersection of significant voxels (p_FDR<0.05) of the results given in Figure 4A-B as a mask. Then a small volume correction (SVC) was applied to 6mm radius spherical regions of interest (ROIs) comparing the attention relative to non-attention to the auditory task. The results are shown in Figure 5 and Table 2 with considerable activity (T=3.17) in left inferior frontal gyrus (−45,24,24; p_FDR<0.044), left superior temporal gyrus (−57,51,6; p_FDR<0.018 SVC corrected) and right superior temporal gyrus (57,−33,3; p_FDR<0.028 SVC corrected). The SVC analyses are based on coordinates given in previous studies of attentional demands (Zhang et al. [48] [−42,13,20]; Kiehl et al. [37] [−62,−34,10]; Zatorre et al. [49] [58,−33,11]). These regions are consistent with sites reported in the literature as reflecting auditory attentional demands. The IFG is considered to be involved with pitch change detection [50,51] and the superior temporal gyrus is a brain region that have been shown to be active in studies investigating auditory short-term functional plasticity [52]. Although our results show stronger hemodynamic responses during the attended condition, Jäncke et al. [52] found a decrease of activation during the course of a 1-week training session. As they reported, one of the reasons for this contradiction might be due to differences with respect to the duration and type of stimulation. While they compare “before” vs. “after” training findings we focus on the responses “during” training. We also analyzed the condition when subject is paying attention to the visual stimuli. Activity in occipital region (Table 3) is higher during attended visual trials (Figure 6) than during attended auditory trials (Figure 5). Previous imaging data have demonstrated that focusing attention on stimuli in one sensory modality increases activity in cortical regions that process stimuli in the attended modality [36,53,54]. Given the lack of any behavioral learning effect it is unlikely that the visual stimuli would evoke a visual learning response. Because of that this paper concerns attention to auditory stimuli only.

Since we were interested in assessing learning performance we used the subject’s specific performance gain over each session in the design matrix. The difference between final and initial thresholds was used as regressors in the general linear model for the auditory attended condition. For the second level analysis, intersubject performance differences were accounted for using the overall performance gain as weights in the design matrix. The results are shown in Figure 7 and Table 4 (uncorrected p<0.005). With this
procedure we could assess the areas involved in learning as the behavioral data was used as regressors in the data estimation. Small volume correction was performed in the same regions as in Figure 5 with a VOI (volume of interest) of 6mm radius. fMRI activity (T=3.23) were observed in left frontal (-45,15,36; pFDR<0.002; SVC corrected), left temporal (-57,-51,24; pFDR<0.002; SVC corrected) and right temporal (60,-39,15; pFDR<0.001; SVC corrected). The substrates underlying rapid learning-induced changes in the auditory cortex are not yet known but they appear to be concerned with perception and selective attention.

EEG data
Figure 8 shows time frequency plots of scalp site Cz for auditory stimulation and Oz for visual stimulation.

EEG and fMRI
Current dipoles were selected within a radius of 6mm from the estimated current peak in each ROI reported in the fMRI analysis (left frontal [IFG: -45,15,36], left temporal [LSTG: -57,-51,24] and right temporal [RSTG: 60,-39,15]). Time frequency analyses were carried out using event-related spectral perturbation ERSP (EEGLAB, [55]) over each of these current dipoles. In this procedure, EEG power within identified frequency bands is displayed relative to power of the baseline period EEG. Blocks of auditory deviant relative to blocks of visual deviant were used to investigate neuronal oscillation at each region of interest. The time-frequency analysis over each current dipole at these areas reveals a different pattern of activation for each subject. Figure 9 shows the
statistical results of the attention versus non-attention condition at regions IFG, LSTG and RSTG over activity localized on the cortex as well as at electrodes F7, T7 and T8 for scalp data. The $t$-statistics of all 11 subjects is performed against null hypothesis of zero mean ($p<0.05$). It can be seen that the responses in LSTG span a wider range compared to the RSTG response, which is more localized in frequency (10 to 20Hz: alpha and beta ranges). The IFG response peaks at around 200ms, later than the temporal cortices as would have been expected. The different responses of neuronal structures in the brain that are frequency band specific have been discussed in the literature in terms of event-related synchronization and desynchronization (ERS/ERD). Quantification of ERS/ERD in time and space has been extensively investigated showing that these responses are functionally related to cognitive processing [56-60]. In this work peak current amplitudes from each region of interest were averaged regardless of phase. This procedure enhanced stimulus-related EEG changes both phase-locked (i.e. event-related potentials) and non-phase-locked (i.e. event-related synchronization and desynchronization) to stimulus onset. Table 5 shows the correlation between EEG power at each frequency band and behavioral threshold at each region of interest (IFG, LSTG and RSTG). Statistical $t$-tests were carried against the hypothesis of null mean at each frequency band. Significant activity were found in IFG at low gamma range ($p<0.05$ corrected) and marginally non significant in RSTG at beta ($p=0.07$ corrected) and low gamma ($p=0.06$ corrected) ranges. Just for comparison learning analysis was conducted with data at scalp sites F7, T7 and T8 (located above the IFG, LSTG and RSTG respectively). Time-frequency plots of scalp data are shown in Figure 9. Although it is inaccurate to assume that the sensor over an area is mainly reflecting activity just below it we tested the correlation between the energy of each frequency range and behavioral data (Table 6). After correcting for multiple comparisons no significant thresholds are found for the different channels. As can be seen by comparison with the activity source localized to the surface of the cortex there are differences in the mixed activity recorded at the electrodes and the cortical activity in the brain region underneath the electrode.

**Table 1 Activated areas during auditory and visual stimulation. MNI coordinates of peak activity of clusters (pFDR<0.05)**

| Brain region | MNI coordinate |
|--------------|----------------|
| Auditory vs. rest | Temporal | −48 -3 -27 |
|                |            | 48 12 -30 |
|                | Frontal   | −42 24 39 |
|                |            | 53 33 0 |
|                | Parietal   | −39 -33 42 |
| Visual vs. rest | Occipital | −39 66 6 |
|                |            | −42. -72 0 |
|                |            | 27 -87 -6 |
|                |            | 39. -63 6 |
|                | Temporal   | −39 -36 15 |
|                |            | 36 -33 6 |
|                | Parietal   | 30 -33 39 |

**Table 2 Attentional effect: MNI coordinates of peak activity clusters (T=3.17)**

| Brain region | MNI coordinate |
|--------------|----------------|
| Temporal lobe/sub gyral B21 | −39. -6,-15 |
| SFG B6   | −9, 12, 66 |
|          | −9, 3, 63 |
|          | −48, 27, 24 |
| MFG B16  | −57, 12, 24 |
| IFG B45  | 57, -33, 3 |
| MTG B22  | −57, -51, 6 |
| MTG B22  |                  |

**Discussion**

The results obtained in this study suggest that attention can be involved and contribute to rapid improvements in specific brain activity during short periods of training.
Both behavioral and physiological data indicate significant activity for attention specific to auditory task within frontal and temporal areas. We suggest that one component of rapid learning is modulated by selective attention, as evidenced by the engagement with the specific task. Our results fall into the category of early attention theories that support that sensory information being used for processing is modified by attention while non-attended features are discarded [1].

Earlier studies of selective attention [37,61] have shown attention-related enhancements of several auditory evoked electromagnetic signals with early modulation at 20-50ms after stimulus onset. The neural source of this early modulated component has been localized in the posterior part of the superior temporal gyrus. The finding of increased responses to attended auditory stimuli suggests the existence of rapid cortical plasticity. Alain et al. [29] have shown that minutes of classical conditioning are sufficient to induce changes of neural responses and receptive field properties in auditory cortices. This plasticity has also been demonstrated by [62] during an experiment of deafferentation of the adult auditory cortex. Their results show a reorganization of cortical representations occurred within a time period of a few hours. In our work, with approximately 80 minutes of training, an improvement in auditory frequency perception could be observed as the subject’s threshold decreased. These results support the theory that during perceptual learning, a fast improvement, occurring early in training, can be induced by a limited number of trials if specific sensory input is provided.

Auditory selective attention
The main result of the beta and gamma oscillations found in the study of the correlation between behavioral thresholds and the energy of the current peak values for each trial suggests that plasticity is also manifested as an increase in the power of induced beta and gamma band activity (GBA, >30Hz) in IFG and RSTG (Table 5). The present correlation pattern in IFG and RSTG during attention demands is consistent with findings of gamma band induction during selective attention [63,64]. However, no significant correlation was found for the LSTG. Although GBA enhancements have been reported in multisensory integration [65], selective attention [66] and memory [67] the way these oscillatory synchronizations are involved with cognitive representations is still not fully understood. The reasons for the presence of activity at and before time zero are unclear. One hypothesis of this early response is that it can be a consequence of some form of anticipatory processing [68]. Alternatively it may be a result of the fast stimuli presentation paradigm. At short ISIs the ERP responses to successive stimuli may overlap, distorting the ERP averages. The activity before time zero can, therefore, be a response to previous stimulation. This explanation has been claimed by some researchers to be more plausible than the occurrence of anticipatory phenomena [69].

Moreover, the finding of task related increased activity in frontal and temporal areas is consistent with the hypothesis that the frontal area is involved with prediction and top-down modulation of auditory selective attention that gives rise to auditory perceptual learning. Our current finding of activity in the superior temporal cortices are in accordance with studies that reported enhanced effects of auditory attention in higher association areas when one modality is attended and the other is ignored [36]. Since attentional effects are very dependent on the task, the

| Table 3 MNI coordinates of peak activity clusters of visual attention (T=3.11) |
|---------------------------------|----------------|
| Brain region                    | MNI coordinate |
| Occipital lobe/ITG              | −48, -69, 0    |
| Temporal lobe/ Fusiform gyrus BA37 | 42, -57, -12  |
| Occipital lobe/ MOG BA19        | −30, -87, 15   |

| Table 4 Learning effect: MNI coordinates of peak activity clusters (T=3.23) |
|---------------------------------|----------------|
| Brain region                    | MNI coordinate |
| Parietal lobe/ postcentral gyrus | −48, -18, 51   |
|      | −48, -30, 57   |
| Temporal lobe/ supramarginal gyrus BA40 | −60, -48, 21 |
| STG B22                         | −54, -54, 9    |
| MFG B9                          | −48, 18, 33    |
|      | −48, 15, 24    |
|      | 45, -39, 3     |
| Temporal lobe/ sub-gyral B22    | 60, -39, 15    |
| STG 22                          |                |
Figure 8 Time frequency representation at Cz for auditory stimulation and Oz for visual stimulation.

Figure 9 Statistic tests (p<0.05) carried out on the time-frequency representation of current dipoles in the 3 ROIs analyzed for the auditory versus visual condition. t-test over time-frequency bins of 11 subjects (10 degrees of freedom). Time frequency analysis was done over activity localized on the cortex in b) IFG, d) RSTG and f) LSTG as well as over channel level activity in a) F7, c) T8 and e) T7. In red: bins whose statistics are greater than the null hypothesis of zero mean. In blue: bins whose statistics are smaller than the null hypothesis of zero mean.
behavioral data was checked. Significant values are marked with (*).}

Table 5 Mean and standard error of correlation coefficients between Fourier transformed source activity and behavioral threshold values for each subject

|          | Alpha         | Beta          | Low gamma    | Gamma        |
|----------|---------------|---------------|--------------|--------------|
| IFG      | (0.06, SE=0.08) | (0.17, SE=0.08) | (0.23, SE=0.07) | (0.05, SE=0.12) |
| RSTG     | (0.07, SE=0.08) | (0.20, SE=0.08) | (0.33, SE=0.07) | (0.15, SE=0.13) |
| LSTG     | (0.07, SE=0.07) | (0.05, SE=0.11) | (0.09, SE=0.08) | (0.04, SE=0.11) |

Energy value for each band was computed and correlation between behavioral data was checked. Significant values are marked with (*).

Table 6 Mean and standard error of correlation coefficients between Fourier transformed scalp activity and behavioral threshold values for each subject

|          | Alpha         | Beta          | Low gamma    | Gamma        |
|----------|---------------|---------------|--------------|--------------|
| F7       | (0.08, SE=0.11) | (0.12, SE=0.11) | (0.25, SE=0.12) | (0.05, SE=0.12) |
| T8       | (0.02, SE=0.09) | (0.14, SE=0.11) | (0.21, SE=0.12) | (0.15, SE=0.13) |
| T7       | (0.04, SE=0.09) | (0.08, SE=0.11) | (0.22, SE=0.09) | (0.15, SE=0.13) |

Energy value for each band was computed and correlation between behavioral data was checked. Significant values are marked with (*).
Control conditions
The STG and IFG have been implicated in several functions beyond that of auditory processing including speech and language processing [84] and social cognition [85]. Our experimental paradigm was carefully designed to account selectively for attention and learning in response to the stimuli presented. To avoid potential confounds caused by anticipation effects the presentation order of the stimuli was randomized. In addition, the time between stimulus presentations was also randomized. To reduce the effects of acoustic noise contamination produced by the fMRI scanning procedure on the cognitive state of the subject we used a sparse presentation procedure in which stimuli were presented in silent periods between scans. To eliminate any biasing effects the same number of deviants and standards were used in the EEG analysis as well as the fMRI analysis. The stimuli themselves did not contain any specific speech, linguistic, or emotion related information that may produce activity in the regions found in our experiment.

In experiments with visual stimulation unconscious involuntary eye movement may be present. These micro-saccades are related to visual fixation and have been shown to have crucial influence on analysis and perception of the visual environment. They can also give rise to EMG eye muscle spikes that can distort the spectrum of the scalp EEG and mimic increases in gamma band power [86]. Some researchers have explored the modulation of synchronous activity by micro-saccades within the primate visual pathway. Yuval-Greenberg et al. [87] have recently noted that spikes in gamma-band activity have a large amount of variability from trial to trial and much of the activity is centered near the eyes. However their results also show a correlation between the amount of gamma band activity and coherence of the image that is shown. In their experiment, during incoherent images micro-saccades were less evident than when the images have some meaning. Melloni et al. [88], however, suggest that saccade related activity is not necessarily trivial and can be related to important cognitive processes that precede, coincide or follow micro-saccades. Recent reports have shown a link between micro-saccades and cognitive processes such as attention, which is not surprising as there is an overlap between the neural systems contributing to control of attention and control of eye movement. There has been a consensus that micro-saccade rates are modulated by both endogenous and exogenous attentional shifts [89]. Additionally, results reporting microsaccades gamma induced activity as being predominantly distributed over the occipital and central scalp [90]. Our results are found in frontal and temporal areas and are not time locked to the onset of the visual stimuli as the control condition was presented randomly.

The source estimation algorithm
In this work we demonstrated the variational hierarchical Bayesian method proposed by Sato et al. [47] applied to EEG data. The hierarchical variational Bayesian method is a source estimation algorithm that incorporates functional magnetic resonance imaging (fMRI) activity as a hierarchical prior [47,91]. It also incorporates structural MRI data to obtain subject specific information about the position and orientation of the current dipoles. The fMRI information determines the prior distribution of the variance in the cortical current. In the hierarchical Bayesian method, the variance of the cortical current at each source location is considered an unknown parameter and is estimated from the EEG signal by introducing a hierarchical prior on the current variance. Although the first papers with VBMEG demonstrated its applications to MEG data [47,91,92] recent papers have been published since then showing that this technique is appropriate to EEG as well [93]. Aihara et al. [94] applied VBMEG to EEG data by incorporating near-infrared spectroscopy (NIRS) as a hierarchical. VBMEG is, therefore, a multimodal encephalography estimation method.

In this experiment we used VBMEG to get better spatiotemporal resolution that is able to extract localized learning related activity that is mixed at level of sensors. As shown in Table 6 this information can not be obtained from activity recorded at the electrodes as it is inaccurate to assume that the activity at a specific sensor reflects the brain activity just underneath it [95-97].

Conclusion
The current study explores the advantage of simultaneous fMRI and EEG recording to investigate brain activity during rapid perceptual learning. Behavioral results suggest that listeners can improve quickly at identifying deviant from standard tones. Rapid improvement in task performance is accompanied by plastic changes in the sensory cortex as well as superior areas gated by selective attention. Moreover, the correlation between ERP time-frequency response and results from behavioral test gives support to our hypothesis of learning during short training periods.

Competing interest
The authors declare that they have no competing interests.

Authors’ contributions
ACS participated in the design of the study, performed the experiments, processed and analyzed data, collected and drafted the manuscript. HCY contributed to the analyses of the results obtained and revised the manuscript critically. MS provided recommendations for the experimental design as well as the EEG and fMRI simultaneous analysis. DC conceived the study, participated in its design and coordination and edited the manuscript. All authors read and approved the final draft.
Acknowledgments
We thank CNPq, CAPES and FAPEMIG for the financial support. We also thank Taku Yoshioka and Ryosuke Hayashi for assistance with the EEG analysis; Yamato Shimosaka, Ichiro Fujimoto and Sayoko Takano for technical support at the Brain Activity Imaging Center at ATR. Akiko Callon for the assistance with the experimental paradigm programming and Yuka Furukawa for the assistance in running the experiments. This research was supported in part by a contract with the National Institute of Information and Communications Technology entitled, ‘Multimodal Integration for Brain Imaging Measurements’.

Author details
1Universidade Federal de São João del-Rei, Ouro Branco, Brazil. 2Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. 3ATR Neural Information Analysis Laboratories, Kyoto, Japan.

Received: 13 August 2012 Accepted: 26 December 2012

References
1. Jagadeesh B, Selzer M, Clarke SE, Cohen LG, Duncan PW, Gage FH: Different dynamics of performance and brain activation in the time course of perceptual learning. Neuroimage 2010, 48(2):697–807.
2. Callan DE, Tajima K, Callan AM, Kubo R, Masaki S, Akahane-Yamada R: Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. Neuroimage 2003, 19(1):113–124.
3. Demany L: Perceptual learning in frequency discrimination. J Acoust Soc Am 1985, 78(3):1119–1120.
4. Ben-David BM, Campeanu S, Tremblay KL, Alain C: Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. Psychophysiology 2010, 48(6):797–807.
5. Moore DR, Amitay S, Hawkey D: Auditory perceptual learning. Learn Mem 2003, 10:88–85.
6. Ahissar M, Cantero JL, Domínguez-Marín E: The time course of neural changes underlying auditory perceptual learning. Learn Mem 2002, 9:138–150.
7. Wright B, Sabin A: Perceptual learning: how much daily training is enough? Exp Brain Res 2007, 180:727–736.
8. Alain C, Snyder JS, He Y, Reineke KS: Changes in auditory cortex parallel rapid perceptual learning. Cereb Cortex 2007, 17:1074–1084.
9. Fukuyama T, Yuki N, Osada S: Auditory perceptual learning: a review. J Exp Psychol 2009, 138:770–782.
10. Callan DE, Callan AM, Kroos C, Varlakots-Bateson E: Multimodual contribution to speech perception revealed by independent component analysis: a study of continuous and discontinuous speech signals. J Cogn Neurosci 2003, 15(8):1197–1209.
11. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. 422.
12. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
13. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
14. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
15. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
16. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
17. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
18. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
19. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
20. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
21. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
22. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
23. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
24. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
25. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
26. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
27. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
28. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
29. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
30. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
31. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
32. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
33. Alain C, Arnott SR, Hevenor S, Graham S, Grady CL: Attentional control of early perceptual learning. J Cogn Neurosci 2003, 15(8):1197–1209.
