Behavioral/Systems/Cognitive

Auditory Cortical Plasticity in Learning to Discriminate Modulation Rate

Virginie van Wassenhove[1] **and Srikantan S. Nagarajan**[2]

[1]Division of Biology, California Institute of Technology, Pasadena, California 91125, and [2]Department of Radiology, University of California at San Francisco, San Francisco, California 94143

The discrimination of temporal information in acoustic inputs is a crucial aspect of auditory perception, yet very few studies have focused on auditory perceptual learning of timing properties and associated plasticity in adult auditory cortex. Here, we trained participants on a temporal discrimination task. The main task used a base stimulus (four tones separated by intervals of 200 ms) that had to be distinguished from a target stimulus (four tones with intervals down to ~180 ms). We show that participants' auditory temporal sensitivity improves with a short amount of training (3 d, 1 h/d). Learning to discriminate temporal modulation rates was accompanied by a systematic amplitude increase of the early auditory evoked responses to trained stimuli, as measured by magnetoencephalography. Additionally, learning and auditory cortex plasticity partially generalized to interval discrimination but not to frequency discrimination. Auditory cortex plasticity associated with short-term perceptual learning was manifested as an enhancement of auditory cortical responses to trained acoustic features only in the trained task. Plasticity was also manifested as induced non-phase–locked high gamma-band power increases in inferior frontal cortex during performance in the trained task. Functional plasticity in auditory cortex is here interpreted as the product of bottom-up and top-down modulations.

Key words: **auditory cortex; evoked magnetic fields; gamma band; plasticity; learning; time; MEG; synchronization**

Introduction

The inherent dynamics of neural computations in the adult brain provide a fascinating example by which an anatomically stable system maintains functional plasticity throughout life. Although the hardwiring of the adult brain limits the extent of functional restructuring (Sereno, 2005), the nervous system is highly versatile at various structural and functional levels (i.e., from neural receptive field to dynamics of neural populations) across a wide range of temporal scales (Gilbert, 1998). The goals of the current study were to examine the early stages of long-term auditory perceptual learning in a modulation-rate discrimination task. Although previous studies on auditory learning in the temporal dimension have used intensive training (as much as 10 d), the first 3 d of training leads to learning that has not yet reached asymptote (Wright et al., 1997; Karmarkar and Buonomano, 2003). Here, we thus hypothesized that the underlying neural mechanisms for the early stage of learning would reflect rapid plasticity in auditory cortex.

Perceptual improvements most likely involve profound restructuring of the underlying neural substrates associated with the long-term representation of a novel (but learned) stimulus together with the increased sensitivity to the trained stimulus set (Schoups et al., 2001; Ghose, 2004). In auditory perceptual learning, the trained perceptual dimension and the task difficulty determine the efficacy, rate, and specificity of learning. For instance, training on a pitch/frequency discrimination task leads to rapid and robust learning (Demany, 1985; Irvine et al., 1985; Moore et al., 2003; Hawkey et al., 2004), although optimal learning may depend on task difficulty (Amitay et al., 2006). In contrast, recent studies show that the discrimination of time intervals is enhanced in the auditory and somatosensory modalities only after intensive training (Wright et al., 1997; Nagarajan et al., 1998; Karmarkar and Buonomano, 2003). To date, very few studies have addressed the neural mechanisms underlying auditory perceptual learning, and virtually none have specifically focused on the time dimension despite its fundamental role in the perceptual categorization of acoustic events. Likewise, the effect of task, attention level, and stimulus context on learning-induced auditory cortex plasticity has primarily been overlooked, despite recent evidence that top-down effects result in important modulatory effects in primary sensory areas (Li et al., 2004; Petkov et al., 2004).

Here, we thus ask whether 3 d of training suffice to significantly improve the discrimination of temporally modulated tone trains and whether the hypothesized improvements are associated with systematic plasticity in auditory cortex. All experiments (training, pretraining, and posttraining sessions) (see Table 1) were performed while participants were recorded under magnetoencephalography (MEG). In the pretraining (day 1) and posttraining (day 5) sessions, task and attention-related effects were tested by examining responses to untrained tasks and to passive

Received July 26, 2006; revised Dec. 20, 2006; accepted Jan. 8, 2007.

This work was supported by National Institutes of Health Grant R01 DC004855 (S.S.N.). We thank members of the Biomagnetic Imaging Laboratory, especially Susanne Honma and Anne Findlay for their excellent technical support. We thank Drs. Merav Ahissar, Dean Buonomano, Christoph Schreiner, and Dan Polley for insightful comments on a previous version of this manuscript. We thank the anonymous reviewers for their astute and helpful comments on this manuscript.

Correspondence should be addressed to Dr. Srikantan Nagarajan, Biomagnetic Imaging Laboratory, Department of Radiology, Box 0628, 513 Parnassus Avenue, S362, San Francisco, CA 94143-0628. E-mail: sri@radiology.ucsf.edu.

DOI:10.1523/JNEUROSCI.4844-06.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/272663-10$15.00/0

listening of learned and novel modulation rates. Both learning of modulation-rate discrimination and the generalizability of learning were tested by examining transfer to the frequency (pitch) and the time (interval) dimensions (for a description and examples of training and test stimuli, see supplemental data, available at www.jneurosci.org as supplemental material).

Materials and Methods

Participants

Nine volunteers (six females; mean age, 31 years) took part in the study. All were strongly right-handed according to the Edinburgh handedness questionnaire (Oddfield, 1971) and had normal hearing and normal or corrected-to-normal vision. Volunteers were paid for their participation. The experiment was conducted in accordance with the Committee on Human Research of the University of California at San Francisco and the Declaration of Helsinki.

Stimuli

In all experiments, the smallest component of the stimuli was a 30-ms-duration 1 kHz tone pip (including a linear 5 ms rise/fall.) A schematic description of all stimuli is provided in the supplemental data (available at www.jneurosci.org as supplemental material). In the temporal modulation rate discrimination task (trained task), the BASE stimulus consisted of four such tone pips modulated at 5 Hz. Precisely, the "temporal modulation rate" here refers to the rate at which the tones were being presented. The standard stimulus or BASE consisted of a 5 Hz temporal modulation tone train, i.e., four tone pips with 200 ms stimulus onset asynchrony (SOA). Each trial consisted of a BASE and one of five TARGET stimuli. The TARGETS consisted of a similar four tone train but this time modulated at variable rates of 5.03125, 5.0625, 5.125, 5.25, or 5.5 Hz (i.e., from SOA of ~198 to ~181 ms). BASE and TARGET stimuli were presented with an average interstimulus interval of 1.5 s.

Additional passive and active controls were used (see below, Procedure) in the pretraining and posttraining sessions (day 1 and day 5.) Active controls were two-interval, two alternative forced-choice tasks, hence identical to the trained task. Each trial consisted of a pair of stimuli to be compared (i.e., the BASE and one of five TARGET stimuli). The first control task consisted of an interval discrimination paradigm, in which all stimuli had the same temporal specifications as in the trained task but differed in the number of tones composing the stimuli; namely, the stimuli consisted of two tones instead of four tones. In this control, the BASE consisted of two tone pips modulated at 5 Hz, and the TARGETS were modulated at 5.03125, 5.0625, 5.125, 5.25, or 5.5 Hz. In a second control task, a frequency discrimination paradigm was used, in which the BASE stimulus was identical to the trained task (a tone train composed of four tone pips modulated at 5 Hz) whereas the TARGETS consisted of a 5 Hz modulated tone train with variable carrier frequencies of the tone pip, which could take any of the following values: 1.0025, 1.005, 1.01, 1.015, or 1.025 kHz.

Two additional controls were tested under passive listening conditions, in which participants attentively listened to the stimuli without requirements to discriminate. The first control or "localizer" consisted of the presentation of a single 1 kHz tone pip at a rate of 0.5 Hz (i.e., well below the rate of the tone train stimuli used in previous tasks). This localizer control was presented before and after each MEG session on each day. The second passive control consisted of the presentations of 5 and 13 Hz modulated tone trains that were pseudorandomly intermixed with an average interstimulus interval of 1.5 s. The 5 Hz modulated tone train was identical to the BASE used in the trained task. The 13 Hz modulated tone train solely differed from the BASE in its modulation rate (13 Hz instead of 5 Hz, otherwise composed of four 1 kHz tone pips).

Procedure

The study took place over 5 consecutive days at the same time of day (i.e., morning, afternoon, or evening). Training days 2–4 consisted only of a temporal modulation rate discrimination task for which participants received feedback after each trial (as well as a passive localizer control.) During each 1 h daily training session, participants were tested on four blocks of 200 trials for a total of 2400 trials in the 3 d of training (800 trials per day and per 1 h of MEG experiment.) Hence, participants received feedback 480 times for each possible BASE/TARGET discrimination.

Table 1. Study timeline within and across days

Day 1	Days 2–4	Day 5
Pretraining	Training	Posttraining
No feedback	Feedback	No feedback
Initial localizer	Initial localizer	Initial localizer
Passive listening	Rate discrimination	Rate discrimination
Frequency discrimination	Final localizer	Interval discrimination
Interval discrimination		Passive listening
Rate discrimination		Frequency discrimination
Final localizer		Final localizer

Each task was no more than 14 min long and was tested while participants were recorded with MEG. The order of the tasks was identical as depicted for all participants.

In the pretraining (day 1) and posttraining (day 5) sessions, several controls were included together with the trained task for a total of ~1.25 h of MEG recording. The order of task presentation was identical for each participant, as described in Table 1. First, the trained temporal modulation rate discrimination task was tested this time without feedback. This presentation block consisted of 200 trials (40 presentations of each possible BASE/TARGET discrimination). Additionally, frequency and interval discrimination tasks were introduced, and each of them consisted of the presentation of 200 trials (40 presentations for each possible BASE/TARGET discrimination). No feedback was provided on any of these tasks. Two additional passive listening controls were also tested during these sessions. In the passive conditions, participants were told to be attentive to the stimuli, although no response was required. First, 5 and 13 Hz temporally modulated tone trains were presented 100 times each. This block lasted only ~5 min for a total of 200 stimuli. Note that the 5 Hz temporally modulated tone train was entirely identical to the BASE of the trained task. Second, the localizer controls consisted of presenting 100 times a 1 kHz tone pip (simplest constituent of the stimuli used in the trained and control tasks), before ("initial") and after ("final") each MEG recording sessions. This localizer control was tested every day (i.e., days 1–5), before and after a session.

In all active tasks, each trial consisted of a pair of stimuli (one TARGET and one BASE) presented in random order (i.e., the BASE could be followed by the TARGET or the TARGET could be followed by the BASE). The tasks were a two-interval alternative forced choice, and participants were asked to report which stimulus in the pair was the "slowest" (or the "lowest" in the frequency discrimination task) by pressing one of two buttons labeled "first" and "second." During the training sessions, participants received feedback after each response. The feedback was provided visually on a monitor screen and consisted of a white symbolic pattern ("####") that changed color according to the correctness of the response; when participants' discrimination between the BASE and the TARGET was correct, the pattern turned green, and, when it was incorrect, the pattern turned red. The feedback cue lasted 300 ms. In the passive listening controls and in the "no feedback sessions" (pretraining and posttraining days), the same visual #### pattern was constantly displayed on the screen, always in white color. The interstimulus intervals were pseudorandomized and partially determined by the reaction time of the participants. In all sessions, the auditory stimuli were presented dichotically through Etymotic earphones (Etymotic Research, Elk Grove Village, IL). The sound level was set to a comfortable hearing level of ~75 dB.

MEG and magnetic resonance imaging recordings

Data were collected using a whole-head MEG system (275 channels; VSM MedTech, Port Coquitlam, British Columbia, Canada) at a sampling rate of 1200 Hz. The position of the head in the MEG Dewar relative to the MEG sensors was determined before and after each block by means of three small coils placed at landmark sites (i.e., nasion and left and right preauricular). The MEG localization information was used in conjunction with magnetic resonance imaging (MRI) scans [1.5 T scanner (GE Healthcare, Milwaukee, WI); flip angle, 40°; repetition time, 27 ms; echo time, 6 ms; field of view, 240 × 240 mm; 1.5 mm slice thickness; 256 ×

256 × 124 pixels]. The three coils were put at the same locations for both the MEG and MRI scans, and the coregistration of the MRI data and the MEG localization results was based on these three fiducial points. The coregistration was used to produce the final image showing both anatomical and functional information. Note that, because of technical and personal reasons, we obtained an MRI scan in only four of the nine participants.

Data analysis

Psychophysics. An individual's performance in each task was computed on the basis of correct responses. Psychometric functions were constructed for each presentation block and for each session. Psychophysical data were fitted to psychometric curves using a maximum-likelihood methodology via "Psignifit" routines (Wichmann and Hill, 2001). The resulting fits allowed estimates of each individual's perceptual thresholds.

MEG-evoked responses. The analysis of evoked responses obtained with MEG was made using commercial software provided by the MEG manufacturer. The average evoked magnetic fields were obtained by averaging the epochs obtained with each stimulus presentation in each block and for each order of presentation on an individual basis. Epochs contaminated by muscle or blink artifacts over 1.5 pT (10% slope) were disregarded. Individuals' average data were then bandpass filtered at 2–40 Hz. Classic auditory evoked magnetic field (AEMF) (m50, m100, and m200) were parameterized in time (latency, in milliseconds) and amplitude [root mean square (RMS)] for each individual. Resulting parameterization was then submitted to statistical analysis.

Statistics. Repeated-measures ANOVAs were performed using SPSS software (SPSS, Chicago, IL). Specific factors and parameters for psychophysical and MEG data are detailed in Results.

MEG dipole fits. Dipole fits were computed using commercial software provided by the MEG manufacturer. Equivalent current dipoles (ECD) of average MEG responses were obtained by assuming a spherical conducting medium with an origin based on individual MRI scans. The localizer data were used to determine the location of the m100 responses in left and right auditory cortices. The resulting position and orientation of the ECD were fixed (after nonsignificant changes between blocks and sessions; for details, see Results and Fig. 4), and a spatiotemporal fit was then performed for the remaining of the average responses in the other tasks. Only sources with a goodness of fit higher than 85% were accepted. The dipole moments (Q value) were analyzed for each test (temporal modulation rate, interval, frequency discrimination, and passive listening). Dipole fits were performed on all nine participants.

Adaptive spatial filtering analysis. This analysis was only performed on four of nine participants, i.e., participants for whom MRI scans were available and for the trained task data (rate discrimination task). Adaptive spatial filtering was used to estimate sources activity modulation within specific frequency bands. To perform this procedure, single-trial MEG data (up to 100 trials per participant and condition) were divided into segments corresponding to baseline (prestimulus) and active periods (250 and 500 ms, respectively, in sliding windows of 250 ms with 125 ms overlap). An estimate of the source activity at each voxel in the brain, based on the MEG data, was calculated as $S_{\mathbf{r}}(t) = \boldsymbol{w}_{\mathbf{r}}^T \boldsymbol{m}(t)$, where $\mathbf{r}$ is the position of the voxel, $S_{\mathbf{r}}(t)$ is the strength of the dipole moment at location $\mathbf{r}$, $\boldsymbol{w}_{\mathbf{r}}^T$ is a vector of spatial filtering coefficient that operates on the data, and $\mathbf{m}(t)$ is the data vector of magnetic field measurements at time t. Therefore, an estimate of source power at each voxel in the brain is given by $S_{\mathbf{r}}^2(t) = [\boldsymbol{w}_{\mathbf{r}}^T \boldsymbol{m}(t)]^2$, and integrating over time yields source variance $S_{\mathbf{r}}^2 = \boldsymbol{w}_{\mathbf{r}}^T \mathbf{C} \boldsymbol{w}_{\mathbf{r}}$, where the covariance matrix is $\mathrm{C} = \mathrm{MM}^{\mathrm{T}}/T$, with M being the spatiotemporal data matrix. Solving for $\mathbf{w}$ by minimizing source variance, subject to $\boldsymbol{w}_{\mathbf{r}}^T b_{\mathbf{r}} = 1$, yields,

$$\mathbf{w}_{\mathbf{r}} = \frac{\mathbf{C}^{-1}\mathbf{b}_{\mathbf{r}}}{\mathbf{b}_{\mathbf{r}}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{b}_{\mathbf{r}}}$$

where $b_{\mathbf{r}}$ is the forward solution for a unit current dipole at position $\mathbf{r}$.

The estimate of the power (P) of the activity at each voxel is then given by the following:

$$\mathrm{P} = \frac{1}{\mathrm{b_r C^{-1} b_r}}.$$

The forward solution was computed assuming a multiple local-sphere spherical volume conductor model based on each participant's MRI.

A functional image was then made up of ratio of the power in each voxel reconstructed for the active and control/baseline windows. For each voxel, a pseudo-F value was computed as follows: let $F = A/C$, where A is the active-state source power, and C is the control-state source power at a particular voxel. The pseudo-F value was $pF = F - 1$ if $F > 1$, and $pF = 1 - 1/F$ if $F < 1$. For our subjects, images were computed for a region of interest (ROI) of $x = [-10.0, 10.0\ \mathrm{cm}]$, $y = [-9.0, 9.0\ \mathrm{cm}]$, and $z = [0.0, 14\ \mathrm{cm}]$, relative to the head frame, in 5 mm steps; this ROI enclosed the entire head, in all cases.

Adaptive spatial filtering was performed for each classic frequency range, namely theta (4–7 Hz), alpha (8–12 Hz), beta (12–25 Hz), low gamma (25–50 Hz), and high gamma (70–100 Hz). In Results, the anatomical landmark provided results from localization of the statistically significant power source maps parameterization (Brodmann's nomenclature) given Montreal Neurologic Institute (MNI) coordinates of individuals' normalized MRI (Dalal et al., 2004).

The time course and amplitude of the estimated activity in each voxel was saved for each stimulus condition (BASE vs TARGET) and for each session (pretraining and posttraining). Training effects were quantified by subtracting the pretraining reconstruction image from the posttraining reconstructions, and residual activity was observed to arise from auditory cortex and its immediate environs.

Results

Rapid auditory perceptual learning

First, we show that perceptual learning can be observed in a temporal modulation rate discrimination task with only 3 d of training (i.e., 3 h of training spread over 3 d, 1 h of training per day). Figure 1 shows the average psychometric functions (left row) collected before and after training (open and filled symbols, respectively). Figure 1*a* (left) shows a significant perceptual improvement in the trained task (temporal modulation rate). Individual performances in pretraining and posttraining sessions were submitted to a two-way repeated-measures ANOVA with factors of day (two: pretraining and posttraining) and target rate (five). Significant main effects of training ($F_{(1,8)} = 11.314$; $p \leq 0.01$) and target rate ($F_{(4,32)} = 27.873$; $p \leq 0.001$) were obtained. Although no significant interaction between training and target rate was found, *post hoc* paired t tests between pretraining and posttraining performances showed a pattern of improvement in discriminability consistent with the level of performance achieved on the pretraining session. Specifically, no significant effects were observed for the most difficult TARGETS, i.e., targets 1 and 2 (5.03125 and 5.0625 Hz, respectively) as opposed to targets 3 (5.125 Hz; $p \leq 0.04$) and 4 (5.25 Hz; $p \leq 0.03$). Target 5 (5.5 Hz) showed a marginally significant improvement ($p \leq 0.08$), which is consistent with a high initial discrimination performance (>90%) on this TARGET (Fig. 1*a*, left). Hence, targets 3 and 4 showed most perceptual improvement on this task.

Similar two-way repeated-measures ANOVAs were performed on the untrained control tasks tested in the pretraining and posttraining sessions. In the interval discrimination task (Fig. 1*b*, left), marginally significant improvements were observed before and after training ($F_{(1,8)} = 4.069$; $p \leq 0.078$), along with a significant effect of target rates ($F_{(4,32)} = 17.723$; $p \leq 0.004$), suggesting a partial transfer of learning to the interval discrimination task. No significant interactions of training with target rates was obtained, but *post hoc t* tests between pretraining and

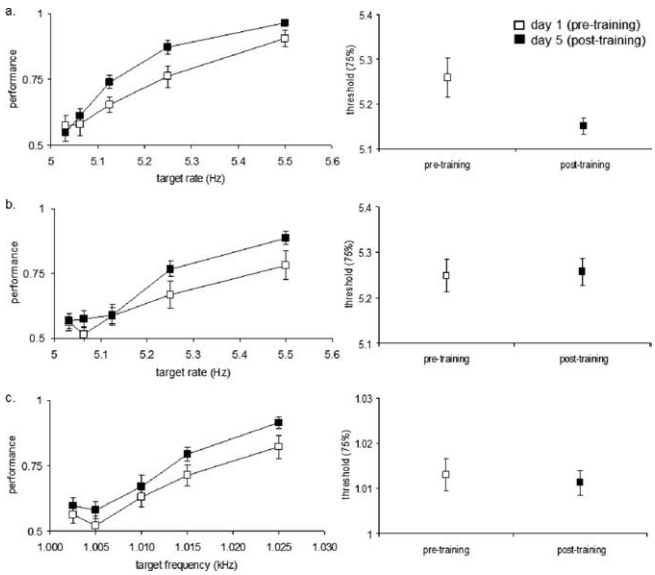

Figure 1. Psychometric curves and perceptual thresholds before and after training. Open symbols are pretraining performances (day 1), and filled symbols are posttraining performances (day 5). The bars correspond to the SEM across all participants ($n = 9$). Perceptual improvements for the trained task (***a***, left) consist of an enhanced performance and decreased perceptual threshold (***a***, right). In the untrained interval discrimination task (***b***), a perceptual improvement was observed (left) but without significant threshold variation (right); in the frequency-discrimination task (***c***), no significant difference was observed between the pretraining and posttraining sessions with respect to performance (left) or perceptual threshold (right). Psychophysical results suggest a partial transfer of learning to the trained perceptual dimension (time) but not to the untrained dimension (frequency/pitch).

posttraining performances revealed marginally significant learning effects on target 4 ($p \leq 0.06$) and target 5 ($p \leq 0.07$). These results suggest a partial transfer of learning consistent with the improvements observed in the trained task. To the contrary, no learning was observed in the control frequency discrimination task ($F_{(1,6)} = 0.278$; $p \leq 0.617$), suggesting that the transfer of learning is limited to the temporal dimension.

To address the pretraining and posttraining threshold variations, individual psychometric functions were fitted using PsigniFit (version 2.5.6) (Wichmann and Hill, 2001). Figure 1 (right) shows the average perceptual thresholds (75%) derived from the Psignifit Weibull fits. Individuals' perceptual thresholds were submitted to one-way repeated-measures ANOVA with factor of day (two: pretraining and posttraining). A significant decrease of the perceptual threshold was observed in the trained rate discrimination task ($F_{(1,8)} = 7.039$; $p \leq 0.029$) (Fig. 1*a*, right). However, neither the interval discrimination task ($F_{(1,8)} = 0.045$; $p \leq 0.838$) (Fig. 1*b*, right) nor the frequency discrimination task ($F_{(1,8)} = 0.275$; $p \leq 0.614$) (Fig. 1*c*, right) showed a significant lowering of the perceptual threshold.

Together, the psychophysical results suggest that 3 d (i.e., 3 h) of training suffice to refine auditory sensitivity to temporal modulation rates. Learning is specific in that only the trained task leads to a significant lowering of the perceptual threshold. Nevertheless, the learning partially transfers to some temporal intervals (i.e., same perceptual dimension) but not at all to the frequency dimension.

Plasticity to learned acoustic feature

The observed learning effects in the trained task were hypothesized to involve changes in low-level acoustic representations, which early auditory evoked magnetic fields are likely to reflect. Figure 2*a* depicts an individual's typical averaged magnetic field traces observed in the left (blue) and right (green) hemispheric sensors to the presentation of the BASE stimulus. The classic m100 and m200 evoked magnetic fields correspond to a temporally localized increased of RMS (compared with prestimulus baseline) at ~100 and ~200 ms poststimulus onset, respectively. The amplitude of the m100 and m200 elicited by the first tone of the BASE was quantified as RMS across MEG sensors located near the temporal lobes and independently for each individual. Figure 2*b* reports the quantification of the m100 and m200 RMS before and after training on the trained task for both hemispheres and averaged across all individuals ($n = 9$). Individuals' RMS values were submitted to a three-way repeated-measures ANOVA with factors of day (two: pretraining and posttraining), order (two: BASE followed by TARGET or TARGET followed by BASE), and hemisphere (two: left and right). A significant increase of RMS between the pretraining and posttraining sessions was found bilaterally for the m100 ($F_{(1,8)} = 47.863$; $p \leq 0.0001$) and the m200 ($F_{(1,8)} = 21.553$; $p \leq 0.002$).

Three-way repeated-measures ANOVAs with RMS as dependent variable and factors of day (two), order (two), and hemisphere (two) were then conducted in the untrained tasks and passive listening conditions to determine whether the observed RMS increase in the trained task was specific to the learning experience.

In the untrained interval discrimination task, a significant bilateral effect of training was observed in the m100 ($F_{(1,8)} = 10.697$; $p \leq 0.011$) and in the m200 ($F_{(1,8)} = 59.88$; $p \leq 0.0001$). Although no significant perceptual threshold decrease was observed in this task, several (easy) TARGETS nevertheless showed significant improvements together with a main effect of target rate. These results suggest that the increased RMS observed for the first tone of the BASE stimulus may be related to the trained perceptual dimension in both the temporal modulation rate and interval discrimination tasks. In the untrained frequency discrimination task, no significant changes of RMS were observed in the m100 ($F_{(1,8)} = 0.42$; $p \leq 0.535$), but a significant increase of RMS was observed in the m200 ($F_{(1,8)} = 7.934$; $p \leq 0.023$). These results further suggest that the m100 RMS increase may be more specific to the perceptual dimension than the m200 RMS variation.

To allow direct comparison between the trained task and the control tasks, relative changes in RMS before and after training were indexed on an individual basis as the ratio of posttraining over pretraining RMS for the m100 and m200, for each hemisphere, each task, and on a per individual basis. The RMS always corresponds to the quantification of the response to the first tone in the BASE. Figure 2*c* reports the grand average posttraining over pretraining RMS ratio for the m100 (top) and m200 (bottom) obtained in all experimental conditions. A positive value indicates an RMS increase from the pretraining to the posttraining sessions, whereas a negative value indicates an RMS decrease. For instance, the trained task (leftmost bars) shows a bilateral positive RMS gain for both the m100 and the m200, consistent with the statistical analysis reported above. These results support the hypothesis that perceptual improvements observed in the temporal rate discrimination task are reflected as amplitude increase of the early AEMF.

The RMS gains were submitted to a four-way repeated-measures ANOVA with factors of task (three: temporal modulation rate, interval, and frequency discrimination tasks), order (two), hemisphere (two), and evoked magnetic response (two: m100 and m200). This analysis revealed a significant effect of task ($F_{(2,16)} = 24.24$; $p \leq 0.0001$) and a significant two-way interac-

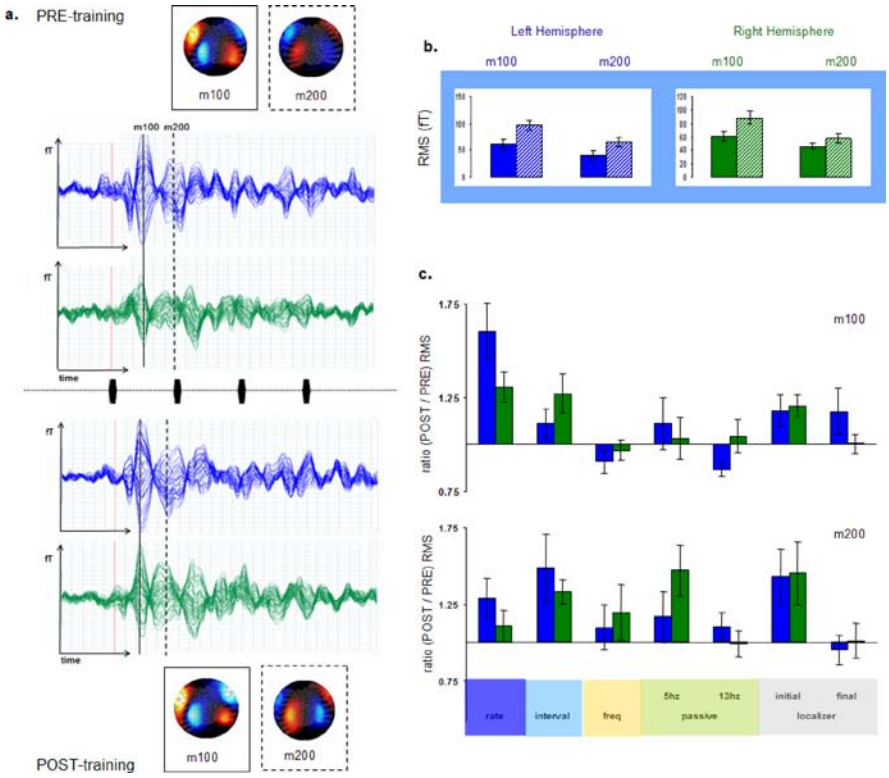

Figure 2. Quantification in sensor space of evoked magnetic fields. ***a***, Typical evoked magnetic fields obtained for one participant in response to the BASE stimulus in the trained task, before (top) and after (bottom) training. Isocontour map distributions of the m100 (line) and m200 (dotted line) are also provided. Left-hemispheric sensors are blue, and right hemispheric sensors are green. ***b***, Grand average RMS of the m100 (left) and m200 (right) before (filled bars) and after (striped bars) training for the BASE stimulus in the temporal rate modulation task (i.e., trained task). ***c***, Grand average ratio of the posttraining RMS over the pretraining RMS for the m100 (top) and m200 (bottom) and for each hemisphere (left sensors are blue, and right sensors are green). Results from the trained task (***b***, ***c***) show a significant bilateral increase of RMS for both m100 and m200. For detailed RMS analysis in control tasks, see Results.

tion of task with hemisphere ($F_{(2,16)} = 6.364$; $p \leq 0.009$); a two-way interaction of task with evoked magnetic response ($F_{(2,16)} = 15.197$; $p \leq 0.0001$) and a three-way interaction of task with evoked magnetic response and hemisphere ($F_{(1,8)} = 7.462$; $p \leq 0.026$) were observed. *Post hoc* paired *t* test comparisons between the RMS gains observed in the trained and in the interval discrimination tasks showed a significant difference in the left-hemispheric m100 gain ($p \leq 0.006$) and in the right-hemispheric m200 gain ($p \leq 0.025$).

Although RMS increase is observed in both trained and control tasks (in particular, the interval discrimination task), the degree to which the evoked responses RMS increases from pretraining to posttraining differs. The m100 and m200 relative RMS increase suggests that the evoked responses are differentially sensitive to stimulus context and task demands. It is important to note that, in all of these conditions, responses to the same physical stimulus (namely, the first tone of the BASE) were analyzed, yet clear differences between the posttraining versus pretraining sessions were observed. In particular, the m100 responses appear more sensitive to the trained perceptual dimension (time) than the m200 responses that are enhanced in all conditions. Additionally, a hemispheric differentiation is noticeable in the frequency discrimination task (frequency/pitch perceptual dimension), in which only the right-hemispheric m200 shows an enhanced response to the BASE stimulus. The trained perceptual dimension is thus particularly important in the changes characterizing the m100. Specifically, a major difference between the processing of the BASE in the temporal modulation rate (trained) and the interval discrimination tasks versus the frequency discrimination task resides in the "nature of the acoustic feature to be extracted," namely the temporal and frequency relationships, respectively. The m100 RMS plasticity shows sensitivity to this difference: although the physical properties of the stimuli are identical, the m100 RMS gain between pretraining and posttraining sessions nevertheless differs. To the contrary, the m200 RMS gain is significant regardless of the task requirements.

To further test this feature-specificity hypothesis, we looked at passive listening conditions tested before and after training on the temporal rate modulation discrimination task. The task attentional demands (trained and untrained tasks) are here compared with passive listening conditions in which no feature is expected to be preferentially extracted. Additionally, to test whether training on a ~5 Hz temporal modulation rate would generalize to a different temporal resolution, a novel temporal modulation rate (13 Hz) was introduced.

The passive listening of a 5 Hz modulation rate (i.e., the BASE) was first considered. Three-way repeated-measures ANOVA on the RMS in response to the first tone of the BASE was performed with factors of AEMF (two), hemisphere (two), and day (two). No significant increase of RMS was found in the left hemisphere (m100: $F_{(1,8)} = 0.223$, $p \leq 0.649$; m200: $F_{(1,8)} = 0.057$, $p \leq 0.818$). In the right hemisphere, marginally significant RMS increases were observed only for the m200 (m100: $F_{(1,8)} = 0.061$, $p \leq 0.811$; m200: $F_{(1,8)} = 6.732$, $p \leq 0.032$). These results are in line with the hypothesis that the early m100 RMS plasticity reflects task-dependent feature extraction, whereas the later right-hemispheric m200 RMS increase is less specific (also observed in the frequency-discrimination task). A similar analysis performed on the responses to the presentation of a novel 13 Hz temporal modulation rate revealed no significant changes of RMS (left hemisphere, m100: $F_{(1,8)} = 0.346$, $p \leq 0.573$; m200: $F_{(1,8)} = 0.770$, $p \leq 0.406$; right hemisphere, m100: $F_{(1,8)} = 0.090$, $p \leq 0.772$; m200: $F_{(1,8)} = 0.002$, $p \leq 0.968$). This result suggest that the early stage of learning is rate specific and does not readily generalize to a different temporal resolution.

In summary, bilateral changes in the m100 response are sensitive to the kind of acoustic features to be extracted with regards to the task demand; a right-hemispheric m200 increase appears to reflect a more automatic but attention-dependent analysis of the learned BASE stimulus.

Posttraining gain in RMS correlates with learning

To further evaluate the specificity of the RMS increase, an analysis of the auditory evoked magnetic fields elicited by the presentation of the TARGET was performed with the hypothesis that the posttraining increase in RMS could either be associated with the learning of the BASE or with an improved sensitivity to the trained stimulus set.

An analysis of the magnetic fields evoked by the presentation of the TARGET stimuli was performed on the pretraining and posttraining datasets. Twenty presentations of each TARGET stimulus were available, and a classic evoked response quantification such as the one performed on the BASE stimulus was not possible because of insufficient signal-to-noise ratio. Nevertheless, six of nine participants provided sufficiently high signal-to-noise ratio to permit an RMS analysis based on temporal windowing. Individual averages of the 20 trials per TARGET were performed, and the RMS was quantified in nonoverlapping bins of 250ms. Figure 3*a* shows a typical example of such an average for one individual. Each window was then submitted to three-way repeated-measures ANOVA with factors of day (two), target (five), and hemisphere (two). A marginally significant two-way interaction of days with target was obtained for the second and third windows ($F_{(4,6)} = 2.292$, $p \leq 0.095$; and $F_{(4,6)} = 3.515$, $p \leq 0.025$, respectively). Note that the second and third windows are, this time, the response intervals to the presentation of the second and third tones in a given TARGET stimulus. The ratio of posttraining over pretraining performance and RMS was computed, providing a relative measure of perceptual improvement and RMS signal change of MEG recordings. Figure 3*b* reports the performance ratio or normalized learning (red) together with the left (blue) and right (green) hemispheric RMS ratio as a function of TARGET rate (target rates are reported on the abscissa). A high degree of correlation was found between the gain in performance (learning ratio) and the RMS gain obtained to the presentation of the fourth TARGET (Fig. 3*c*). This correlation was mainly observed in the left hemisphere. No other significant correlations between RMS change and performance were obtained, and these data are thus not reported.

Is there a generalized cortical sensitivity enhancement?

Before and after each daily session, a localizer consisting of a single 1 kHz tone was used to evaluate the participant's head position relative to the MEG sensor array within and across each day (see above, MEG dipole fits). This condition also served as baseline for possible changes of RMS during passive listening of the smallest feature (1 kHz tone) characterizing the BASE or TARGET stimuli, which were otherwise composed of four such tones (two in the interval discrimination task). Three-way repeated-measures ANOVAs were conducted on the m100 and m200 RMS as dependent variables and with factors of day (two), position within session (two: initial and final), and hemisphere (two). A significant effect of days was observed bilaterally for the localizer recorded in the initial position within sessions (m100: $F_{(1,8)} = 7.739$, $p \leq 0.024$; m200, $F_{(1,8)} = 5.56$, $p \leq 0.046$) but not for the m200 of the localizer tested at the end of the sessions ($F_{(1,8)} = 5.928$, $p \leq 0.041$; m200, $F_{(1,8)} = 1.508$, $p \leq 0.254$). However, *post hoc* analysis showed significant and marginally significant differences between the m100 RMS gain in the trained task and the initial localizer (left hemisphere, $p \leq 0.02$; right hemisphere, $p \leq 0.06$, respectively), as well as the final left-hemisphere localizer ($p \leq 0.02$.)

This result suggests that the sensitivity of the auditory cortex to a 1 kHz tone has been refined during the course of the experiment but that increased RMS observed previously cannot solely be accounted for by such refinement. We thus conducted additional analysis, this time in source space to address (1) the relative contribution of the auditory sources to the evoked responses variations reported here and (2) the possible involvement of nonauditory cortices in modulating the RMS of the auditory magnetic-evoked responses, motivated by the effects of task requirements and potential top-down modulations of auditory cortices during training.

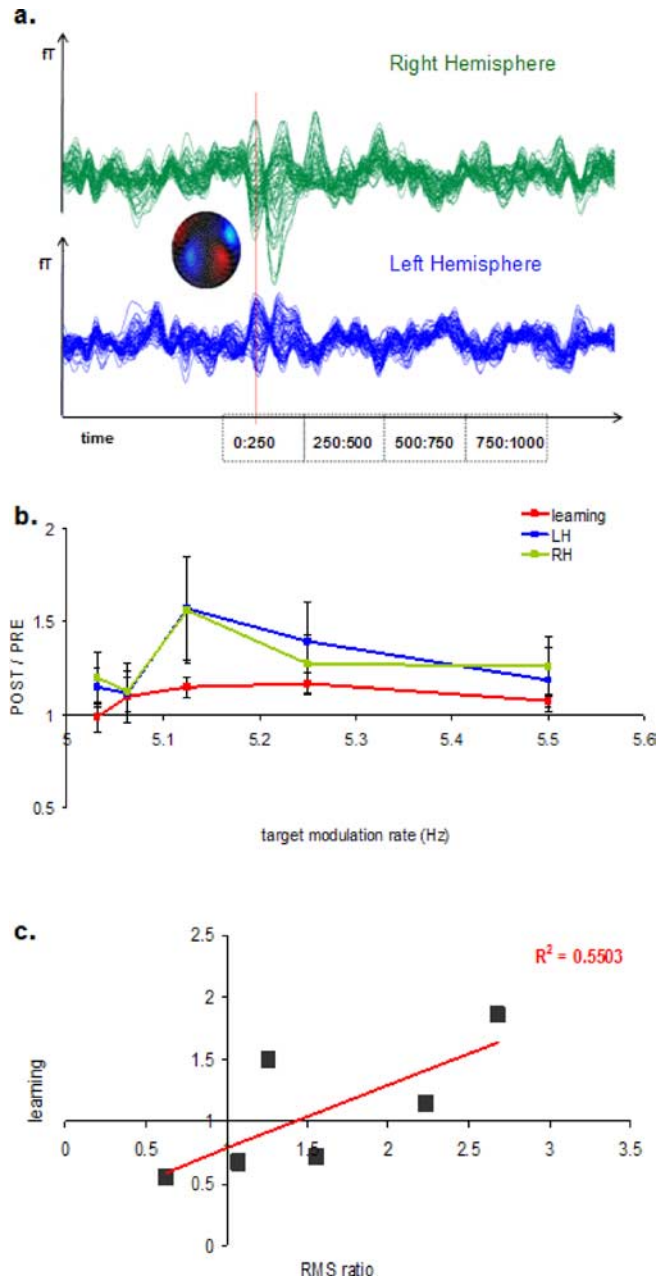

Figure 3. Increased RMS to TARGET stimulus in the temporal rate discrimination task (trained task). The top depicts a typical individual's evoked magnetic fields obtained by averaging 20 trials in response to a TARGET stimulus. Right hemispheric sensors are green, and left-hemispheric sensors are blue. In ***b***, the grand average normalized performance (i.e., the pretraining performance divided by the posttraining performance) is plotted in red as a function of the TARGET rate. Targets 3 and 4 showed the most perceptual improvement. The posttraining RMS observed in the second averaging time bin (i.e., from 250 to 500 ms after stimulus onset) was normalized to the pretraining RMS in the same time bin. Grand averages from the right hemispheric sensors (green) and the left-hemispheric sensors (blue) are plotted as a function of TARGET rate. The gain in RMS is observed mostly for targets 3 and 4. ***c*** reports a positive correlation (0.55) between the "gain in performance" and the "gain in RMS" for the fourth target. Such correlation was not obtained for the third target.

MEG source-space analysis of plasticity for the BASE stimulus

The source reconstruction of the auditory evoked responses was processed in several steps. First, a dipole fit procedure was under-

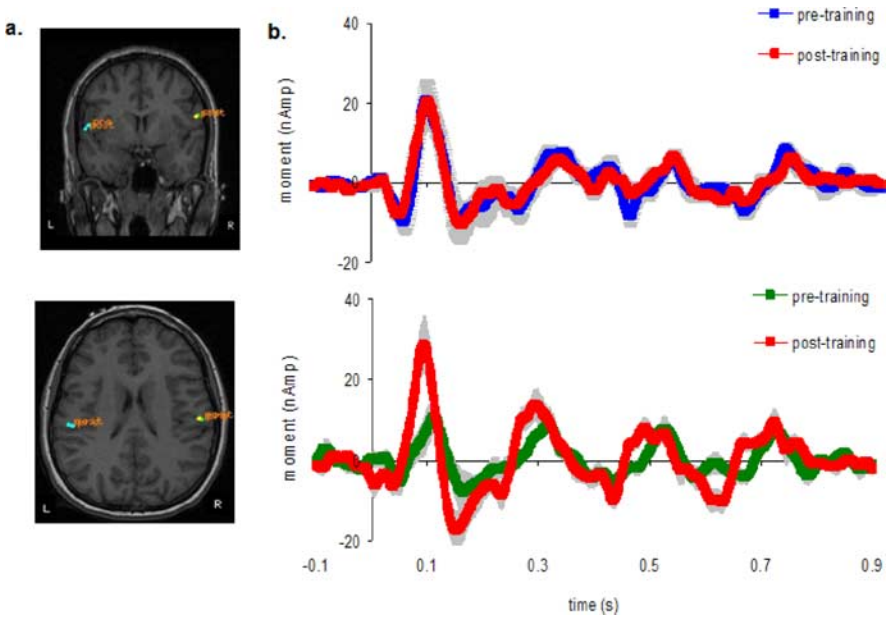

Figure 4. Spatiotemporal dipole fits before and after training. ***a***, Dipole fits to the m100 obtained in response to the presentation of the initial localizer (1 kHz tone presented at the beginning of the MEG session) obtained in the pretraining and posttraining sessions. No significant difference was found in the location of the dipoles. ***b***, Spatiotemporal dipole moment time series obtained for the left (top) and right (bottom) hemispheres, before (blue and green, respectively) and after (red) training. These spatiotemporal dipole moment time series were averaged across all nine participants, and the shaded gray areas report the SEMs. A significant increase of dipole moment in the posttraining session was observed in the right hemisphere but not in the left hemisphere. The following rate observed in the right hemisphere (i.e., increased dipole moment after the presentation of the tone) is absent in the left hemisphere.

taken using the localizer data for all participants. Because the experimental design required subjects to be positioned in the machine several times, one possible issue in our analysis would be that the position of the head across experimental sessions differed. However, the resulting dipole fits performed on pretraining and posttraining sessions did not show significant spatial shifts within or across sessions (i.e., within and across days, respectively). Figure 4*a* shows an example of the superposed dipole fits in one participant measured during the pretraining and posttraining sessions. Three-way repeated-measures ANOVAs were computed with *x*, *y*, and *z* coordinates as independent variables and with factors of hemisphere (two), day (two), and time of presentation within the session (two: initial and final). No significant interaction was found for hemisphere with day (x, $F_{(4,8)} = 1.029$, $p \leq 0.407$; y, $F_{(4,8)} = 0.789$, $p \leq 0.541$; z, $F_{(4,8)} = 0.436$, $p \leq 0.781$) or hemisphere with time within session (x, $F_{(1,8)} = 0.016$, $p \leq 0.903$; y, $F_{(1,8)} = 1.029$, $p \leq 0.340$; z, $F_{(1,8)} = 0.730$, $p \leq 0.418$), indicating that the positioning of participants' head in the MEG Dewar was consistent during the entire course of the study.

Subsequently, the dipole fit obtained with the localizer data were thus taken as template for dipole fitting the MEG sensor signals originating from the auditory cortices in the other experimental conditions. The spatial coordinates and angle of the dipole were fixed in space, allowing to fit the strength of the dipole to the pretraining and posttraining time series obtained in the average data (such as those depicted in Fig. 2*a*). This analysis was done on a per individual basis. The spatiotemporal dipole fits lead to two time series corresponding to the variation of the dipole moment in time. Each hemisphere was fitted separately. Figure 4*b* shows results of the spatiotemporal dipole moments averaged across all nine participants in the trained condition. For statistical testing, the dipole moment series was averaged in 50 ms time bins around the m100 response to each tone in the BASE (e.g., for the first tone dipole moments were average from 75 to 125 ms after stimulus onset). Four-way repeated-measures ANOVA with factor of time window (four), day (two), hemisphere (two), and order (two) showed a marginally significant effect of day ($F_{(1,8)} = 4.29$; $p \leq 0.072$) and a main effect of window ($F_{(3,24)} = 28.078$; $p \leq 0.0001$) on the dipole moments. Additionally, a marginally significant two-way interaction was observed between hemisphere and day ($F_{(1,8)} = 4.142$; $p \leq 0.076$) and a significant interaction observed between hemisphere and window ($F_{(3,24)} = 4.695$; $p \leq 0.01$).

The hemispheric difference can be seen in Figure 4*b* as an increase of the dipole moment in the right hemisphere but not in the left hemisphere. These results do not appear congruent with the sensor space analysis of the evoked responses, in which bilateral increase of RMS was observed in the trained task. A second analysis in source space was thus conducted using an adaptive spatial filtering method to disambiguate these results. This analysis was more conservative in that it was now solely performed on the four participants for whom MRI could be recorded.

Individual averages were imported in NUTMEG (Dalal et al., 2004). Adaptive spatial filtering was performed independently for each hemisphere (see Materials and Methods). A window of ~60 ms was chosen surrounding the m100 peak (from 90 to 150 ms). Figure 5 depicts the pretraining (top) and posttraining (bottom) reconstructions of source activity that is localized to auditory cortex and its immediate environs. The associated temporal profiles are provided beside each result.

Although the spatiotemporal dipole fits did not reflect the bilateral increase in the AEMF, the adaptive spatial filtering analyses performed on average data corroborate the analysis obtained in sensor space. Note that the activations projected on an individual's brain indicate significant differences between the pretraining and postsession activation (i.e., contrast analysis). These results corroborate the sensor space analysis.

The reconstruction techniques based on stimulus-evoked responses (e.g., dipole moment) only relate to phase-locked components of the cortical responses (i.e., responses that are occurring at almost the same time poststimulus presentations). Thus, these methods neglect potential induced activity, i.e., the cortical responses that are evoked by the stimulus presentation but that do not always occur at the same latency after stimulus onset. The variability in the temporal profile of induced responses is eliminated in the averaging procedure, and induced responses do not appear in average data. Hence, to evaluate whether non-phase-locked cortical sources contribute to the increased RMS of the evoked magnetic fields, additional source-space analysis was performed on the four individuals' single-trial data. Given that the dipole fit does not reveal increased RMS in the left hemisphere, one hypothesis was that the left auditory cortex increased activation could reflect a second source of non-phase-locked activation.

Adaptive spatial filtering analysis

Individuals' MEG data were analyzed in the time-frequency domain using adaptive spatial filtering analysis (see Materials and Methods.) In this technique, a source-space *F* ratio was computed for each time bin, each time frequency, each condition, and each voxel between the prestimulus period and poststimulus period of interest. Results consist in time-frequency spectra in overlapping time bins of 250 ms (125 ms overlap) and frequency bins corresponding to the major and classic functional neuroimaging nomenclature: theta (4–7 Hz), alpha (8–12 Hz), beta1 (12–18 Hz), beta2 (18–25 Hz), low gamma (25–58 Hz), and high gamma (62–98 Hz). Each of the time-frequency spectra provide the peak activations for the entire MRI of the individual with a preselected sensitivity of 1 mm. Consistent results were observed with this method only in the high-frequency gamma band and in three of four participants. Figure 6 shows significant induced activity residuals obtained when contrasting the posttraining versus pretraining sessions (i.e., subtraction of source-space *F* ratio images) for these three individuals. A significant peak of activation in the high-frequency gamma band (62–98 Hz) was observed solely in the left hemisphere and in the inferior frontal cortex (IFC). This result suggests that plasticity may not be confined to auditory cortex but involve other cortices (here, IFC). It is important to note that significant IFC-induced activity observed here occurs within the time window at which we observed increased RMS in the auditory cortices (i.e., within 250 ms after stimulus onset). Together with the absence of increase in the left-hemisphere dipole moment, these results suggest that the IFC activity may contribute to the early plasticity (RMS increase) observed in the left auditory cortices. The right hemisphere did not show significant difference in induced activation. The significant increase of dipole moment in the right hemisphere suggests that plasticity in the right hemisphere may be confined to auditory cortices. No other contrasts lead to significant and consistent results with this particular analysis.

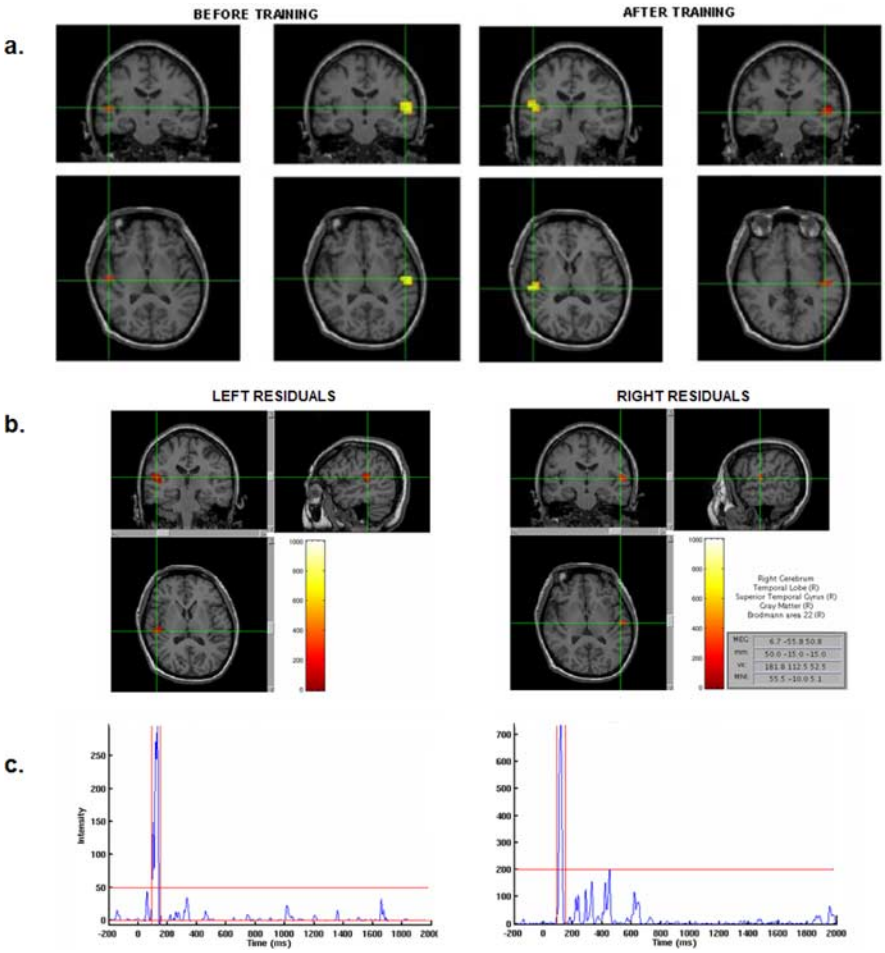

Figure 5. Beam-forming source reconstruction of the auditory cortex response to the BASE stimulus before and after training. ***a***, Sources of activation in a 60 ms window surrounding the m100 peak in the left and right hemispheres, before (left panels) and after (right panels) training. ***b***, Subtracting the pretraining activation from the posttraining activation in source space reveals a positive residual in both hemispheres (left-hemisphere data are in the left section of ***b***, and right-hemisphere data are in the right section of ***b***). This result converges with the RMS results obtained in sensor space (i.e., with the bilateral RMS increase of the auditory evoked responses). Note, however, that this analysis reveals residual activation in the left hemisphere, which was not seen with spatiotemporal dipole fit. ***c***, The time course of the residual activation is confined to the m100 and m200, and the analysis was here limited to the m100 peak. The left (***c***) depicts the time course of activation in the left hemisphere, and the right (***c***) is the time course of activation in the right hemisphere. The red vertical lines delimit the time window used for source reconstruction.

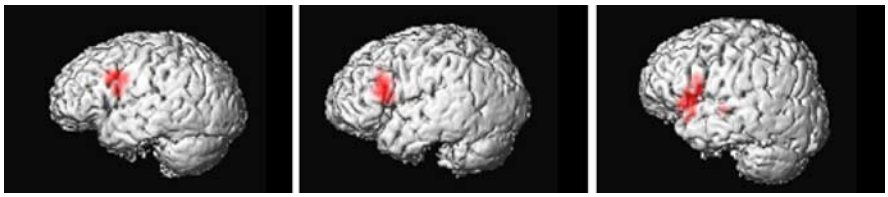

Figure 6. Induced residual high gamma activity (62–98 Hz) in the inferior frontal cortex after training. Time-frequency adaptive spatial filtering analysis was performed on four participants. Three of four participants shown here exhibited a consistent residual peak of activation in the contrast analysis, corresponding to an increased synchronization state in the high gamma band (62–98 Hz) when comparing the posttraining session with the pretraining session activation. Specifically, this residual was localized to the inferior frontal cortex (via MNI coordinate normalization procedure, the peak activation appears more specifically located in the inferior frontal gyrus for each participant).

Discussion

In this study, we have shown that 3 h of training (spread over 3 d) can significantly improve auditory temporal perception. The reported learning remains specific: a partial transfer of learning was found in a temporal interval discrimination task but not in a frequency discrimination task. The observed auditory perceptual learning was accompanied by auditory cortical plasticity manifested as an enhancement of early bilateral auditory evoked-magnetic field responses for trained BASE and TARGET stimuli. In particular, the m100 enhancement was found to be specific to the trained BASE stimulus in the rate and interval discrimination tasks but not in the frequency or passive listening tasks; the m200 enhancement was observed in one or both hemispheres in all

experimental conditions. Additionally, the enhancement of the MEG responses to the presentation of the TARGET stimuli was correlated with behavioral improvements, providing additional evidence that the observed plasticity is specific to perceptual learning. Finally, source localization analyses of oscillatory activity during task performance before and after learning showed that plasticity also manifests as an increase in the power of induced high gamma (62–98 Hz) band activity located in the left IFC. This source of synchronization indicates that learned changes are not confined to auditory cortex and provides a potential neural substrate for top-down modulation of learning-induced plasticity in early sensory cortices.

Previous studies of auditory temporal interval discrimination learning have used intensive training (>10 sessions) (Wright et al., 1997; Karmarkar and Buonomano, 2003). In these studies, learning was found to asymptote after 5–7 d, with maximal learning occurring within the first few days of training. Here, we examined learning with 3 d of training to investigate nonasymptotic properties of plasticity accompanying auditory perceptual learning. Our generalizations are consistent with the previous intensive training studies that have demonstrated transfer across tasks within the same perceptual dimension (Wright et al., 1997). The absence of full transfer and generalization suggests that participants have not yet reached an asymptotical stage of learning, despite a significant improvement in the trained task. The lack of transfer of learning to untrained tasks and conditions is often interpreted as reflecting plasticity at early stages of the sensory processing hierarchy.

Some aspects of the rapid auditory perceptual learning obtained here may be accounted for by an early stage of plasticity in the auditory cortex, observed as an increase of early evoked magnetic responses to the presentation of the BASE, TARGET, and localizers. These effects are consistent with previous MEG and electroencephalographic studies that have shown neural changes characterized by enhanced responses to trained stimuli, presumably resulting from a larger cortical recruitment that occurs across multiple timescales of training ranging from a few minutes to several days (Cansino and Williamson, 1997; Pantev et al., 1999; Menning et al., 2000; Atienza et al., 2002; Bosnyak et al., 2004). The general finding of increased responses to the trained stimulus is also consistent with a long history of neurophysiological studies that have shown the existence of rapid cortical plasticity, in which minutes of classical or operant conditioning suffice to induce profound changes of neural responses and receptive field properties in primary auditory cortices (Galambos et al., 1956; Weinberger and Diamond, 1987; Buonomano and Merzenich, 1998). Neurophysiological research has characterized different kinds of plasticity in the adult auditory cortex (Weinberger, 2004; Ohl and Scheich, 2005). For instance, in conditioning paradigms, auditory neurons selectively increase their responses to conditioned and unconditioned stimuli (Weinberger et al., 1990; Cruikshank and Weinberger, 1996; Polley et al., 2004; Ohl and Scheich, 2005). In learning-induced paradigms, in which an animal is trained to discriminate between stimuli, neurons initially unresponsive to a stimulus set are recruited for the representation of the learned stimuli (Recanzone et al., 1993; Rutkowski and Weinberger, 2005). An increase of the auditory evoked magnetic fields is thus consistent with the notion that the sensitivity of the auditory cortex for a 1 kHz tone has been refined after training; the local feature of the trained stimuli may provide one kind of bottom-up plasticity.

In our study, the BASE was identical in the rate, frequency, and passive tasks, but some differences among the conditions involved the instructional set (whether to attend and which stimulus feature to process) as well as stimulus context. Although stimulus context could have influenced the findings, we feel that influences related to instructional set were more likely to have been important. All stimuli in the training and control tasks shared very close temporal and frequency properties that are likely recruiting a comparable set of neurons within a single auditory neural bandwidth and within a single isofrequency stripe, respectively (Heil et al., 1992; Schreiner, 1995; Schreiner et al., 2000; Barbour and Wang, 2003). Thus, at the level of a tone train, temporal rates may share modulation transfer functions observable within a single neural population (Barbour and Wang, 2002, 2003). However, in studies of visual perceptual learning, the characteristics of generalization from learning have been shown to be modulated by task difficulty. In so doing, plasticity can manifest in "reverse order" to the sensory hierarchy (Ahissar and Hochstein, 2004; Fahle, 2005). Recent psychophysical results point to a similar effect of task difficulty in auditory learning (Amitay et al., 2006), and top-down influences in auditory perceptual learning should thus be considered in accounting for auditory cortex plasticity. Our study explicitly incorporated different tasks and attentional demands, and learning-induced plasticity did show specificity after such factors.

An additional finding is the posttraining residual activation of the left IFC, suggesting that plasticity is not confined to auditory cortices and rather engages a distributed network. This is in agreement with results obtained in monkey neurophysiology using a similar training task (Machens et al., 2005) and in line with the neuroanatomy of the auditory cortex (Romanski et al., 1999; Kaas and Hackett, 2000). One possible interpretation is that the IFC activation reflects the involvement of the working memory system (Gottlieb et al., 1989), enabling feedback of the learned stimulus set on auditory cortices (Pasternak and Greenle, 2005). If such were the case, perceptual categorization in learning does not solely involve the improved representation of the base stimulus (or template) in early sensory cortices but also engages in a discriminative procedure between the BASE and the TARGET stimuli via the interfacing of the working memory system with the sensory cortices. In this context, the increased RMS observed after learning may reflect the template status as direct comparison of the internalized BASE representation and the incoming TARGET. Such hypothesis converges with recent neurophysiological and functional MRI findings that show task-specific plasticity in early sensory cortex (Fritz et al., 2003; Ohl and Scheich, 2005). In effect, early plasticity in auditory cortex may integrate early on feedback signals from working memory and attentional systems with incoming auditory inputs. For instance, auditory cortical responses were shown to be highly dependent on task demands (Fritz et al., 2003), and a recent functional MRI study shows task-attention dependency of auditory cortices (Petkov et al., 2004). Top-down modulations have been recently incorporated in a framework for auditory cortex analysis, in which feedback mechanisms intervene in shaping auditory representations (Scheich et al., 2005). This framework is particularly suitable in light of our results, which suggest that the bottom-up and top-down streams interface in auditory learning and plasticity.

It is noteworthy that auditory plasticity may incorporate an attentional component. The auditory m100 component is particularly sensitive to attention, and plasticity reflected in this early component is seldom observed (Menning et al., 2000) compared with plasticity of the later m200 (Atienza et al., 2002). Attentional selection has been suggested recently to affect neural plasticity in visual perceptual learning (Zoltàn and Sohn, 2005), and this

could possibly be the case in auditory perceptual learning. Here, participants may have learned to pay more attention to the stimuli as part of their learning experience on the trained task. Attention may thus play an important selective role in neuroplastic changes of the auditory response (Menning et al., 2000; Bosnyak et al., 2004).

Our source analyses data suggest a tight following response of the right hemispheric auditory sources, whereas the left-hemispheric sources may engage in a tight coupling with the left IFC. This is suggested by the inadequacy of the dipole fit method to model an increase in the left auditory cortex response: increased activation was found with sensor analysis and with adaptative spatial filtering together with a left IFC activation. This possible functional lateralization is consistent with the recent proposal that the "temporal structure" of acoustic stimuli is bilaterally analyzed through different hemispheric temporal resolutions (Boemio et al., 2005). Several results in our study suggest that the early left-hemispheric responses are particularly tuned to the featural component of the stimuli, whereas the right hemisphere variations shows a more global pattern of response across tasks and stimuli. Together, we interpret our findings as evidence that the state of the auditory cortex in processing learned inputs is fundamentally changed after training. This plasticity ultimately results from the training process, which allowed for the refinement of sensitivity in early auditory cortices, a restructuring partly mediated by contextual demands such as task and attention.

References

Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. Trends Cogn Sci 8:457–464.

Amitay S, Irwin A, Moore DR (2006) Discrimination learning induced by training with identical stimuli. Nat Neurosci 11:1446–1448.

Atienza M, Cantero JL, Dominguez-Martin E (2002) The time course of neural changes underlying auditory perceptual learning. Learn Mem 9:138–150.

Barbour D, Wang X (2002) Temporal coherence sensitivity in auditory cortex. J Neurophysiol 88:2684–2699.

Barbour D, Wang X (2003) Auditory cortical responses elicited in awake primates by random spectrum stimuli. J Neurosci 23:7194–7206.

Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:380–395.

Bosnyak DJ, Eaton RA, Roberts LE (2004) Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40Hz amplitude modulated tones. Cereb Cortex 14:1088–1099.

Buonomano DV, Merzenich MM (1998) Cortical plasticity: from synapses to maps. Annu Rev Neurosci 21:149–186.

Cansino S, Williamson SJ (1997) Neuromagnetic fields reveal cortical plasticity when learning an auditory discrimination task. Brain Res 764:53–66.

Cruikshank S, Weinberger NM (1996) Receptive-field plasticity in the adult auditory cortex induced by Hebbian covariance. J Neurosci 16:861–875.

Dalal S, Zumer J, Agrawal V, Hild K, Sekihara K, Nagarajan S (2004) NUTMEG: a neuromagnetic source reconstruction toolbox. Neurol Clin Neurophysiol 52:1–4.

Demany L (1985) Perceptual learning in frequency discrimination. J Acoust Soc Am 78:1119–1121.

Fahle M (2005) Perceptual learning: specificity versus generalization. Curr Opin Neurobiol 15:154–160.

Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci 6:1216–1223.

Galambos R, Sheatz G, Vernier V (1956) Electrophysiological correlates of a conditioned response in cats. Science 123:376–377.

Ghose GM (2004) Learning in mammalian sensory cortex. Curr Opin Neurobiol 14:513–518.

Gilbert CD (1998) Adult cortical dynamics. Psychol Rev 78:467–485.

Gottlieb Y, Vaadia E, Abeles M (1989) Single unit activity in the auditory cortex of a monkey performing a short term memory task. Exp Brain Res 74:139–148.

Hawkey DJC, Amitay S, Moore DR (2004) Early and rapid perceptual learning. Nat Neurosci 7:1055–1056.

Heil P, Rajan R, Irvine D (1992) Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli. II. Organization of response properties along the "isofrequency" dimension. Hearing Res 63:135–156.

Irvine D, Martin R, Klimkeit E, Smith R (1985) Specificity of perceptual learning in a frequency discrimination task. J Acoust Soc Am 108:2964–2968.

Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA 97:11793–11799.

Karmarkar UR, Buonomano DV (2003) Temporal specificity of perceptual learning in an auditory discrimination task. Learn Mem 10:141–147.

Li W, Piëch V, Gilbert CD (2004) Perceptual learning and top-down influences in primary visual cortex. Nat Neurosci 7:651–657.

Machens C, Romo R, Brody C (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science 307:1121–1124.

Menning H, Roberts LE, Pantev C (2000) Plastic changes in the auditory cortex induced by intensive frequency discrimination training. NeuroReport 11:817–822.

Moore DR, Amitay S, Hanke DJC (2003) Auditory perceptual learning. Learn Mem 10:83–85.

Nagarajan S, Blake DT, Wright BA, Byl N, Merzenich MM (1998) Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. J Neurosci 18:1559–1570.

Oddfield R (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Ohl FW, Scheich H (2005) Learning-induced plasticity in animal and human auditory cortex. Curr Opin Neurobiol 14:470–477.

Pantev C, Wollbrink A, Roberts LE, Engelien A, Lütkenhöner B (1999) Short-term plasticity of the human auditory cortex. Brain Res 842:192–199.

Pasternak T, Greenle M (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6:97–106.

Petkov C, Kang X, Alho K, Bertrand O, Yund E, Woods D (2004) Attentional modulation of human auditory cortex. Nat Neurosci 7:658–663.

Polley D, Heiser M, Blake DT, Schreiner CE, Merzenich MM (2004) Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. Proc Natl Acad Sci USA 101:16351–16356.

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13:87–103.

Romanski L, Tian B, Fritz J, Mishkin M, Goldman-Rakic P, Rauschecker J (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136.

Rutkowski R, Weinberger NM (2005) Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. Proc Natl Acad Sci USA 102:13664–13669.

Scheich H, Ohl FW, Schulze H, Hess A, Brechmann A (2005) What is reflected in auditory cortex activity: properties of sound stimuli or what the brain does with them? In: The auditory cortex. A synthesis of human and animal research (König R, Heil P, Budinger E, Scheich H, eds), pp 383–407. Mahwah, NJ: Erlbaum.

Schoups A, Vogels R, Qian N, Orban G (2001) Practicing orientation identification improves orientation coding in V1 neurons. Nature 412:549–553.

Schreiner CE (1995) Order and disorder in auditory cortical maps. Curr Opin Neurobiol 5:489–496.

Schreiner CE, Read HL, Sutter ML (2000) Modular organization of frequency integration in the primary auditory cortex. Annu Rev Neurosci 23:501–529.

Sereno MI (2005) Plasticity and its limits. Nature 435:288–289.

Weinberger NM (2004) Specific long-term memory traces in primary auditory cortex. Nat Rev Neurosci 5:279–290.

Weinberger NM, Diamond D (1987) Physiological plasticity in auditory cortex: rapid induction by learning. Prog Neurobiol 29:1–55.

Weinberger NM, Ashe J, Diamond DM, Metherate R, McKenna T, Bakin J (1990) Retuning auditory cortex by learning: a preliminary model of receptive field plasticity. Concepts Neurosci 1:91–132.

Wichmann FA, Hill J (2001) The psychometric function. I. Fitting, sampling, and goodness of fit. Percept Psychophys 63:1293–1313.

Wright BA, Buonomano DV, Mahncke HW, Merzenich MM (1997) Learning and generalization of auditory temporal-interval discrimination in humans. J Neurosci 17:3956–3963.

Zoltàn V, Sohn W (2005) Learning to suppress task-irrelevant visual stimuli with attention. Vision Res 25:677–685.