

Distributed Auditory Cortical Representations Are Modified When Non-musicians Are Trained at Pitch Discrimination with 40 Hz Amplitude Modulated Tones

Daniel J. Bosnyak¹, Robert A. Eaton² and Larry E. Roberts¹

¹Department of Psychology, McMaster University, Hamilton, Ontario, Canada L8S 4K1 and ²Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada L8S 4K1

Several functional brain attributes reflecting neocortical activity have been found to be enhanced in musicians compared to non-musicians. Included are the N1m evoked magnetic field, P2 and right-hemispheric N1c auditory evoked potentials, and the source waveform of the magnetically recorded 40 Hz auditory steady state response (SSR). We investigated whether these functional brain attributes measured by EEG are sensitive to neuroplastic remodeling in non-musician subjects. Adult non-musicians were trained for 15 sessions to discriminate small changes in the carrier frequency of 40 Hz amplitude modulated pure tones. P2 and N1c auditory evoked potentials were separated from the SSR by signal processing and found to localize to spatially differentiable sources in the secondary auditory cortex (A2). Training enhanced the P2 bilaterally and the N1c in the right hemisphere where auditory neurons may be specialized for processing of spectral information. The SSR localized to sources in the region of Heschl's gyrus in primary auditory cortex (A1). The amplitude of the SSR (assessed by bivariate T^2 in 100 ms moving windows) was not augmented by training although the phase of the response was modified for the trained stimuli. The P2 and N1c enhancements observed here and reported previously in musicians may reflect new tunings on A2 neurons whose establishment and expression are gated by input converging from other regions of the brain. The SSR localizing to A1 was more resistant to remodeling, suggesting that its amplitude enhancement in musicians may be an intrinsic marker for musical skill or an early experience effect.

Keywords: 40 Hz auditory steady-state response, auditory cortical plasticity, evoked potentials, musical skill, pitch discrimination

Introduction

It is well established that the frequency tuning of neurons in the mammalian auditory cortex is not hardwired after early development but can be altered in the adult brain by experience with behaviorally significant acoustic signals (Buonomano and Merzenich, 1998). Plastic modification induced by aversive conditioning in adult guinea pigs has been documented for neurons in primary (A1, auditory core) and secondary (A2, belt/parabelt) regions of the auditory cortex well as in the medial, dorsal and ventral divisions of the auditory thalamus (Edeline, 1999). When brain regions are contrasted within the same conditioning procedure, tone-evoked plasticity is expressed more commonly by neurons in A2 (96%) than by neurons in A1 (63%; Diamond and Weinberger, 1984). Using owl monkeys, Recanzone *et al.* (1993) found that appetitive discrimination training for small changes in spectral pitch enhanced the cortical territory representing the trained frequencies in A1 by a factor exceeding 5. The sharpness of tuning and temporal response properties of multi-unit recordings were also modified for the trained frequencies

in this study. Training at acoustic discrimination in the owl monkey using amplitude modulated (AM) tones varying either in carrier frequency (Blake *et al.*, 2002) or AM rate (Beitel *et al.*, 2003) increased the spiking activity of A1 neurons for stimuli associated with reward compared to stimuli that were not.

Neural plasticity of the magnitude seen in these animal studies suggests that remodeling of the human auditory cortex by behavioral training should be expressed in auditory evoked potentials (AEPs) and magnetic fields (AEFs) which reflect the activities of populations of neurons in the brain. Consistent with this hypothesis, AEPs and AEFs evoked by musical stimuli are enhanced in musicians who have processed such stimuli extensively in their environment compared to non-musicians who have not. Enhancement has been reported for the magnetic N1m (Pantev *et al.*, 1998), the electrical P2 (Shahin *et al.*, 2003), and the right-sided electrical N1c (Shahin *et al.*, 2003), each of which localizes to spatially differentiable centers of activation in the region of A2 where neuroplastic remodeling is robustly expressed. The auditory N19–P30 middle latency waveform, which has been localized by magnetic and electrical source imaging (Scherg and von Cramon, 1986; Godey *et al.*, 2001; Yvert *et al.*, 2001) and by intracortical measurements (Celesia, 1976; Liégeois-Chauvel *et al.*, 1993; Godey *et al.*, 2001;) to Heschl's gyrus (A1), is also enhanced in musicians (Schneider *et al.*, 2002). This waveform underlies the 40 Hz auditory steady-state response (Galambos *et al.*, 1981; Gutschalk *et al.*, 1999) and is correlated with the anteromedial extent of this anatomic structure and with measured musical skill (Schneider *et al.*, 2002). However, while enhancement of these functional brain attributes in musicians may be of neuroplastic origin, one cannot rule out the possibility that enhancement results from prenatal influences or a genetic code that guides the development of auditory cortex and shapes the decision to train musically.

A more direct approach to assessing the expression of neuroplastic processes in AEPs and AEFs is to measure these responses when subjects are trained at novel acoustic discriminations. Recent studies by Eaton and Roberts (1999), Tremblay *et al.* (2001) and Atienza *et al.* (2002) indicate that at least one transient response of the AEP, the P2 with a latency of ~185 ms, is enhanced when such training is carried out under laboratory conditions. These results are congruent with the hypothesis that enhancement of the P2 AEP when evoked by musical stimuli in musicians (Shahin *et al.*, 2003) is a consequence of the extensive prior experience that musicians have had with such stimuli in the context of musical performance. In the present paper, we evaluated neuroplastic properties of several components of the AEP by training non-musician subjects to discriminate small changes in the carrier frequency of 40 Hz AM pure tones. This stimulus procedure allowed us to

separate the P2 and other transient AEPs of interest whose sources are known to localize to A2 from the 40 Hz steady-state response (SSR) whose cortical sources reside more specifically in A1. Our goals were to (i) determine which of these AEP components reflecting activity in spatially distributed regions of the auditory cortex is sensitive to remodeling by neuroplastic mechanisms, and to (ii) begin to describe the network behavior that underlies remodeling of human auditory cortex by experience.

Materials and Methods

Subjects

Eight subjects (six males) aged 25–30 years participated in 18 sessions of discrimination training and testing. All were graduate students at McMaster University (five right-handed). None had received formal musical training or played a musical instrument. Subjects were paid \$150 for their participation. Subjects gave their written consent following procedures approved by the university ethics committee in conformance with the Declaration of Helsinki.

Training Environment and Auditory Stimuli

All sessions were carried out in an electrically shielded and acoustically dampened room. Auditory stimuli consisted of 10 ms sinusoidal tone pips (onset and offset windowed with a 2 ms cosine² function) of different carrier frequencies presented at 40 Hz for a duration of 1 s. For convenience, we will refer to these stimuli as 40 Hz AM pure tones (see Fig. 1A for the spectrum and time domain waveform of the stimulus at 2 kHz). The stimuli were generated by a Tucker Davis sound generator and delivered through Noisebuster stereo headphones (Noise Cancellation Technologies Inc., Model NB-EX) which actively attenuated background noise by ~15 dB at frequencies below 500 Hz. The intensity of individual 1 s pulse trains was varied randomly between 57 and 60 dB above each subject's measured threshold throughout discrimination training and testing in order to ensure that subjects used pitch and not intensity as the basis for their discriminative choices.

Experimental Procedure

Subjects participated in experimental sessions of the following types, which were given over ~20 days in the order indicated below.

Preliminary Session

The experiment commenced with a preliminary session in which auditory thresholds and frequency discrimination ability were measured using 40 Hz AM tones. Hearing thresholds were determined at 2 kHz for each subject using six cycles of a staircase procedure. Frequency discrimination ability was then evaluated using the staircase method described by Levitt (1971). On each trial, a standard stimulus (S1) and a comparison stimulus (S2) were presented separated by an interval of 0.5 s. Subjects indicated by a button press whether the two tones were of the same frequency (50% of the trials) or different frequencies. Subjects were not informed of the correctness of their decisions. Forty trials were presented for each of nine S1 frequencies between 1.8 kHz and 2.2 kHz in a single test that lasted ~25 min. S2 frequencies differed initially from the S1 by 60 Hz and were adjusted up or down according to the subject's performance. The purpose of discrimination assessment was to select the S2 stimuli that would be used for subsequent discrimination testing and training for each subject. The preliminary session also familiarized subjects with 40 Hz AM tones they would encounter during test and training sessions.

Test Sessions

Two test sessions were administered, one given the day following the preliminary session and the second ~18 days later after the training series (see below) had been completed. Test sessions provided a fine-grained assessment of discrimination ability before and after discrimination training. Each test session consisted of three blocks each containing 360 trials requiring same/different frequency judgments without knowledge of results. The three blocks differed with regard to the set of stimuli used. In one block the standard stimulus (S1) was 2.0 kHz while the comparison frequencies (S2) varied from 2.0 to 2.1 kHz. Because this stimulus set was used later for discrimination training, we refer to it as the 'trained set'. The remaining blocks evaluated 'control' stimulus sets which employed either 1.8 kHz or 2.2 kHz as the S1 stimulus (S2 stimuli were 0–100 Hz higher). The order of assignment of stimulus sets to the three blocks varied between subjects but was the same for each subject before and after training. Control sets allowed us to determine whether changes in behavioral performance and brain activity detected after training

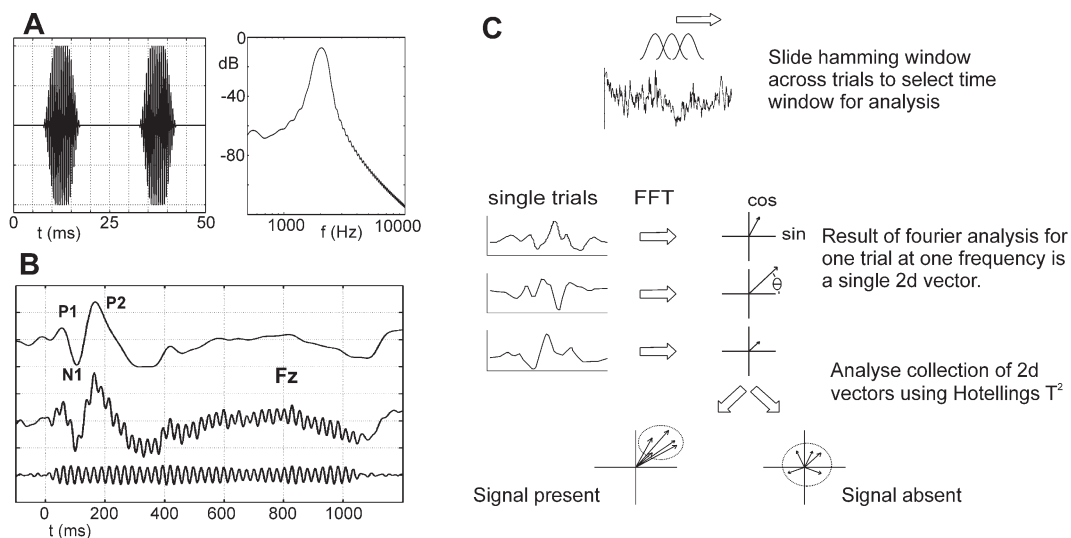


Figure 1. (A) Waveform and spectrum of the 40 Hz AM stimulus at 2.0 kHz (the standard stimulus of the trained stimulus set). (B) Auditory evoked potential elicited by the 2.0 kHz stimulus of A on test trials administered after 15 sessions of training for pitch discrimination. Middle trace: auditory evoked potential (high-pass filtered at 1 Hz) shows the 40 Hz SSR riding on a low frequency transient waveform. Upper trace: P1, N1 and P2 transient responses are set into relief by low-pass filtering at 15 Hz to remove the 40 Hz component. Lower trace: band pass filtering (30–50 Hz) singles out the 40 Hz SSR. Each step on the ordinate is 1 μ V (each trace referenced to zero). (C) Single trial T^2 analysis of the 40 Hz SSR. SSR amplitude on each trial is represented by vector length and phase by the angle θ . Vector end points do not include the origin when a steady state response is present.

related specifically to the trained carrier frequencies, which would be expected if discriminative learning had occurred.

For each test block the procedure was as follows (similar to the Same-Different-Higher procedure described by Jesteadt and Bilger, 1974). On each trial subjects listened to the S1 stimulus followed 0.5 s later by an S2 stimulus. The S2 was either the same frequency as the S1 stimulus (50% of the trials) or one of six different comparison tones which were higher in frequency than S1. The lowest of the six different comparison tones was always 2 Hz higher than the standard, while the highest comparison tone was usually 60 Hz higher than the standard. The four comparison stimuli between these two extremes were chosen by the experimenter on an individual basis such that subjects were likely to detect two of them at least 50% of the time and the other two less than 50% of the time. On each S1/S2 trial the subjects indicated 'same' or 'different' by a button press; the next trial commenced 1 s later. Subjects were instructed to base their response choices on a change in pitch and not stimulus intensity. Test sessions lasted ~90 min (30 min for each block evaluating a single stimulus set) and used identical comparison frequencies before and after training so that performance between the test sessions could be compared.

Training Sessions

Fifteen training sessions were administered which were identical to the test sessions, except for the following differences. Only the stimulus set with the 2.0 kHz S1 was trained. In addition, feedback was given about the correctness of discriminative decisions by two LEDs placed 1 m in front of the subject at eye level. If the subject's response was correct, a green LED was illuminated; if the response was incorrect, a red LED lit up. The LED stayed on for 500 ms. Each training session contained 480 trials and lasted ~30 min. The S2 stimulus was 2.0 kHz on 240 trials ('same') and one of six higher frequencies on the other 240 trials ('different'). Training sessions were scheduled daily with a 1 day pause on the weekend.

An adaptive procedure was applied between the training sessions when the subjects performed without error on more than one comparison frequency. When this happened the highest comparison frequency was removed and replaced with a new comparison frequency in the region where the probability of a correct detection was near 0.5. The comparison frequencies of 2 Hz and 60 Hz were excepted from this procedure and kept constant for all subjects.

Retention Session

The retention session took place 7 weeks after the second test session. The procedure for the retention session was identical to that of a training session. Two of the eight subjects were not available for the retention session.

Analysis of Behavioral Data

Behavioral performance was evaluated at each comparison frequency for each subject. At each comparison frequency which differed from S1 ($\Delta f > 0$) the probability of a 'hit' [$P(H)$] was calculated by dividing the number of 'different' responses (hits) by the number of stimulus presentations. Next, the probability of a 'false alarm' [$P(FA)$] was calculated as the proportion of trials on which the subjects responded 'different' when the S2 frequency equaled the S1 frequency of 2 kHz ($\Delta f = 0$). From these two measures a performance score (P) was calculated for each comparison frequency according to the formula $P = [P(H) - P(FA)] / [1 - P(FA)]$. The measure P corrects $P(H)$ at each comparison frequency for the tendency of the subject to commit false alarms and reaches 1.0 (the maximum value attainable) at Δf s where the subject makes no errors (Green and Swets, 1966). Psychophysical functions were constructed for each subject, test session, and stimulus set by plotting P against all Δf s > 0 and fitting the curve with a logistic. The discrimination threshold was defined as the value of Δf corresponding to $P = 0.5$.

The discrimination performance of each subject was also evaluated by d' using the procedure of Dember and Warm (1979). This metric was calculated for each comparison tone ($\Delta f > 0$) by subtracting a z -score calculated for $P(H)$ from a z -score calculated for $P(FA)$. Values of d' were calculated for training sessions 1-3 grouped together and for training sessions 13-15 grouped together using comparison tones common to both sessions.

Electrophysiological Recording

The electroencephalogram (EEG) was recorded during the two test sessions and during training sessions 3 and 13. For the first two subjects a 19-channel recording was taken (10/20 system, Electrocap, tin electrodes) and for the remaining six subjects a 64-channel recording was made (NeuroMedical QuickCap, Ag/AgCl electrodes). Electrode sites were abraded with a blunt sterile needle and covered with Electro-Gel to lower skin impedance to <10 k Ω . The EEG was sampled at 500 Hz with a DC amplifier (NeuroScan Synamps) and recorded using Cz as the reference electrode and AFz as ground. Data were re-referenced off line to a common average prior to signal processing.

Analysis of EEG Data

EEG data were epoched from 400 ms before stimulus onset to 400 ms after stimulus offset. Epochs were baselined for the interval 50 ms prior to stimulus onset and were linear detrended. Individual epochs were passed through a 60 Hz notch filter and sorted in order of total variance (energy) for artifact rejection. Epochs with the largest variance were rejected until 80% of the trials remained, following the procedure of John *et al.* (2001). When a trial was rejected, the data of the entire epoch were discarded.

EEG responses to the S1 stimulus only were analyzed. These stimuli were processed by the subjects in attention, were presented most frequently in the experiment (360 presentations for the S1s of each stimulus set in the test sessions and 480 presentations of the trained S1 during training sessions), and were uncontaminated by the preparation of behavioral responses. Transient responses of the EEG were analyzed using the 19 electrodes (10/20 array) that were common to every subject ($n = 8$) and recording. Source modeling of the transient responses and analyses of the SSR in 100 ms moving windows (see below) were carried out for the six subjects for whom 64 channel recordings were taken. Signal processing procedures for transient and steady state responses were as follows.

Transient Responses

Epochs were averaged for each subject and EEG recording session. Figure 1B (middle trace) shows this average for the second test session (Fz electrode, high pass filtered at 1 Hz) where a 40 Hz oscillation can be seen to be riding on a slower transient waveform. The averaged data for each subject were filtered 1-15 Hz forward and backward (zero phase shift) with a sixth-order Chebyshev filter to remove the 40 Hz component, thereby setting into relief the transient waveform with prominent P1, N1, and P2 components (upper trace, Fig. 1B). Spherical spline maps of current source density were generated at the amplitude maximum of each component to show scalp topography (64 channel subjects only). The amplitude and latency of P1, N1 and P2 components, and a fourth component (the N1c, not identified in Fig. 1B) which showed properties of interest, were determined by a computer algorithm that searched electrodes containing their amplitude maxima for amplitude peaks occurring within latency windows determined from the grand averaged data. P1 amplitude was recorded as the most positive peak occurring in the Fz electrode between 40 ms and 110 ms after stimulus onset. N1 amplitude was recorded as the most negative peak occurring at Fz between 90 ms and 120 ms after stimulus onset, and P2 amplitude as the most positive peak occurring between 120 ms and 200 ms at this electrode. The N1c was defined as the most negative peak occurring between 120 ms and 180 ms at electrodes T7 and T8 in accordance with the radial orientation of this brain event in each hemisphere (Woods, 1995). The difference between N1 and P2 amplitude (P2 minus N1) was also calculated for each subject. This metric provided a conservative estimate of P2 amplitude by removing possible contributions arising from changes in the overlapping N1 waveform.

40 Hz Steady State Response

The lower trace of Figure 1B shows the time-domain average of the 40 Hz SSR extracted from the trace seen in the middle panel by band pass filtering (30-50 Hz, zero phase shift, eighth-order Butterworth). In principle, changes induced in the 40 Hz SSR by discrimination training could be expressed tonically throughout the S1 stimulus or be confined to more restricted epochs of the stimulation period. In addi-

tion, training could modify the number of neurons activated by the S1 stimulus which would be expected to affect SSR amplitude, or the temporal properties of the neural representation which could influence SSR phase. In order to evaluate these possible effects, a single trial analysis was conducted in which a Fourier transformation was applied within a Hamming window 100 ms wide, zero-extended to 1000 ms, that was moved across the EEG (from 400 ms prior to S1 onset to 400 ms after S1 offset) in 10 ms time steps. Within each window, separately for each subject, trial, and test session, the 40 Hz component was represented as a vector in a polar plot where SSR amplitude was given by vector length and SSR phase by the angle θ in as depicted in Figure 1C. Changes in θ consequent on shifting the window by 10 ms were corrected at each time step. As depicted in the representation of Figure 1C, confidence limits circling the vector endpoints do not include the origin when an SSR is present (see the example of Picton *et al.*, 1987). The likelihood of this outcome under the null hypothesis is distributed as Hotelling's bivariate T^2 (Valdes-Sosa *et al.*, 1987; Victor and Mast, 1991). Rejection of the null hypothesis implies the presence of a 40 Hz SSR of some phase and amplitude in the test session.

This technique gave an assessment of the SSR on each test session administered before and after discrimination training. In order to evaluate before/after training effects on the SSR, we utilized a two-sample version of the T^2 test to contrast the two test sessions (Timm, 1975). This test is conceptually similar to the single sample case and is calculated as:

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{x}_1 - \bar{x}_2)' S^{-1} (\bar{x}_1 - \bar{x}_2)$$

where:

$$S = \frac{A_1 + A_2}{N_1 + N_2 - 2}$$

$$A_1 = \sum_{i=1}^{N_1} (\bar{x}_{1i} - \bar{x}_1)(\bar{x}_{1i} - \bar{x}_1)'$$

$$A_2 = \sum_{i=1}^{N_2} (\bar{x}_{2i} - \bar{x}_2)(\bar{x}_{2i} - \bar{x}_2)'$$

Because the number of T^2 statistics generated for each subject and session was large and not independent for overlapping 100 ms zero-extended windows, critical values of T^2 for statistical evaluation of before/after differences were determined by Monte Carlo simulations conducted separately for each subject as described later.

Significant before/after differences identified by T^2 established that some aspect of the SSR (amplitude, phase, or both) had been modified by training. However, further evaluation was necessary to identify which aspect of the response had changed. For this purpose we calculated for each subject, test session, and 100 ms window the mean phase of the vectors (SSR phase) and mean vector length (the resultant, called herein SSR amplitude) in order to identify which of these measures contributed to before/after differences in the 40 Hz SSR. Also calculated were (i) phase coherence by the method of Picton *et al.* (2001), and (ii) absolute vector length for each test session. Comparison of these two measures between test sessions allowed determination of whether before/after differences in SSR amplitude measured as the mean vector were a consequence of a decrease SSR phase variability around its central tendency, or an overall increase in vector length independent of phase. Absolute vector lengths were normalized with respect to the maximum length observed for each subject before contrasting before/after differences.

Amplitude modulation of the discriminative stimuli permitted identification of a response at the modulation frequency (the 40 Hz SSR) whose cortical sources have been found to localize to the region of Heschl's gyrus in A1. However, transient responses (whose cortical sources are distributed in A2) could in principle contain a 40 Hz spectral component of the transient waveform (for example, the transient gamma band response; Pantev *et al.*, 1991) that is potentially confus-

able with the 40 Hz SSR. To evaluate this possibility, we applied the T^2 method to evaluate 40 Hz activity in a separate control group of eight subjects (undergraduate students paid \$8 per hour) who performed the 2.0 kHz discrimination task for a single session without feedback for correctness, using unmodulated S1 and S2 stimuli. This condition gave an estimate of 40 Hz energy present in transient AEPs evoked by acoustic stimulation when the 40 Hz SSR was absent.

Source Modeling

Source analysis of the average-referenced AEP field patterns (N1, N1c, P2 and the 40 Hz SSR) was carried out using BESA 2000 (MEGIS GmbH, Munich, Germany). Analyses were conducted separately for each stimulus set and test session using the group averaged data. Two regional sources were used to describe the cortical generators for each AEP component (one source in each hemisphere, constrained to localize symmetrically following Scherg and von Cramon, 1986). Sources were determined at the peak of the AEP waveform (root mean squared transformed) within the same latency windows used for analyzing amplitude peaks in the EEG data. Medial/lateral (x), anterior/posterior (y), and inferior/superior (z) coordinates of each regional source were recorded together with dipole moment. The residual variance of the source model averaged 1.4%, 3.5%, 4.8%, and 1.8% for the N1, N1c, P2 and SSR, respectively (2.7% overall), with no fit exceeding 7.0% residual variance. It should be noted that regional sources determined by BESA use three orthogonal vectors (one in each plane) to describe cortical activations contributing to AEPs. These vectors were investigated further as described in the results section, to provide information on the relative contribution of tangential and radial vectors to N1 and N1c transient responses and SSR waveforms.

Statistical Evaluation

Changes in behavioral performance and in transient AEPs induced by discrimination training were evaluated by repeated-measures ANOVAs. Analyses applied to the two test sessions included the variables before/after training and stimulus set (S1 stimuli of the trained set and the two control sets). Pre-planned contrasts were evaluated by conventional *t*-tests and *post hoc* contrasts with the Least Significant Difference test. Peak amplitude and latency were analyzed for the AEPs, and, for behavioral performance, the metrics P , d' , discrimination threshold, and slope of the psychophysical functions determined for each stimulus set. All probabilities are two-tailed unless otherwise stated.

Monte Carlo methods were used to evaluate the 40 Hz SSR. The presence of an SSR for each subject and test session was not in doubt; T^2 for the 40 Hz Fourier component exceeded 45 in all subjects and 100 ms moving windows. In order to contrast the test sessions for training effects, we generated the distribution of T^2 under the null hypothesis for each subject and stimulus set using the procedure of Manly (1991). For each moving window 144 trials were taken at random from the maximum of 288 trials that were available after artifact rejection in the first test session (before training), and a further 144 trials were taken the trials available in the second test session (after training). These 288 trials were used to calculate T^2 when no difference was expected between before/after measurements. This constituted one simulation. One thousand of these simulations were performed for each stimulus set to approximate the distribution of T^2 under the null hypothesis. Although these simulations were conducted separately for each subject, the results across subjects were similar, and we found that a critical value of $T^2 = 8.0$ created a rejection region of $P < 0.01$ for all subjects considered singly. In order to determine a critical value to apply to a T^2 map of a group of subjects, we combined one randomly selected 'null hypothesis' map from each subject into a group mean map, and repeated this process 1000 times to generate a distribution for this map under the null hypothesis. In this case a critical value of $T^2 = 4.5$ was found to depict $P < 0.05$ and $T^2 = 6.0$ to depict $P < 0.01$. The SSR was evaluated at several electrode sites but the response in the 40 Hz region was maximal at Fz and only the results for this electrode are reported.

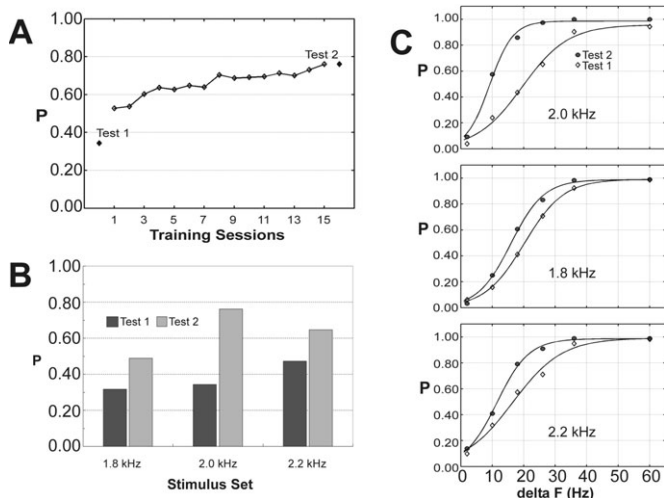


Figure 2. Behavioral performance. (A) The performance measure P (hit rate corrected for false alarms) is plotted over 15 sessions of discrimination training and on the opening (Test 1) and closing (Test 2) test sessions. Data are for the trained 2.0 kHz stimulus set. (B) Performance on test blocks administered before and after discrimination training is contrasted for the trained stimulus set (2.0 kHz) and for control sets (1.8 and 2.2 kHz) above and below the trained stimuli. (C) Psychophysical functions before and after training on the three stimulus sets.

Results

Behavioral Performance

Behavioral performance (P) on the trained stimulus set is shown over the 15 sessions of training in Figure 2A, where performance on the opening and closing test sessions is also depicted. Performance improved rapidly from the opening test session and then more gradually thereafter. A significant main effect of training sessions [$F(14,92) = 4.72, P < 0.001$] was found, as were significant preplanned contrasts between training sessions 1 and 15 [$t(7) = 3.86, P = 0.006$] and 3 and 13 [$t(7) = 2.68, P = 0.03$] which corroborated gradual improvement throughout the training series. Performance on the test sessions given before and after training is contrasted for the trained stimulus set and the two control sets in Figure 2B. Main effects were found for before/after [$F(1,7) = 19.27, P = 0.003$] and for stimulus set [$F(2,14) = 7.32, P = 0.006$] and as well as an interaction of these variables [$F(2,14) = 22.01, P < 0.001$]. Performance improved after training on all three stimulus sets, but more so for the 2.0 kHz set [$t(7) = 6.90, P < 0.001$] than for the 1.8 kHz [$t(7) = 2.66, P = 0.04$] and 2.2 kHz [$t(7) = 2.72, P = 0.03$] untrained stimuli.

Training effects were corroborated by d' and by psychophysical functions calculated for each subject. When averaged over subjects d' increased from 0.99 at the outset of training (sessions 1–3 collapsed) to 1.59 at the end of training (sessions 13–15 collapsed), giving $t(7) = 4.69, P = 0.002$. Psychophysical functions are shown for each stimulus set in Figure 2C. Discrimination thresholds (Δf at $P = 0.5$) decreased from 20.3, 20.2 and 16.7 Hz prior to training for the 1.8, 2.0 and 2.2 kHz sets, respectively, to 9.3 Hz for the trained 2.0 kHz set and to 16.0 and 11.6 Hz for the 1.8 and 2.2 kHz sets, respectively. These results gave rise to a main effect of before/after [$F(1,7) = 7.624, P = 0.028$] and to an interaction with stimulus set [$F(2,14) = 6.289, P = 0.011$] which was attributable to before/after differences appearing for the trained stimuli [$t(7) = 2.99,$

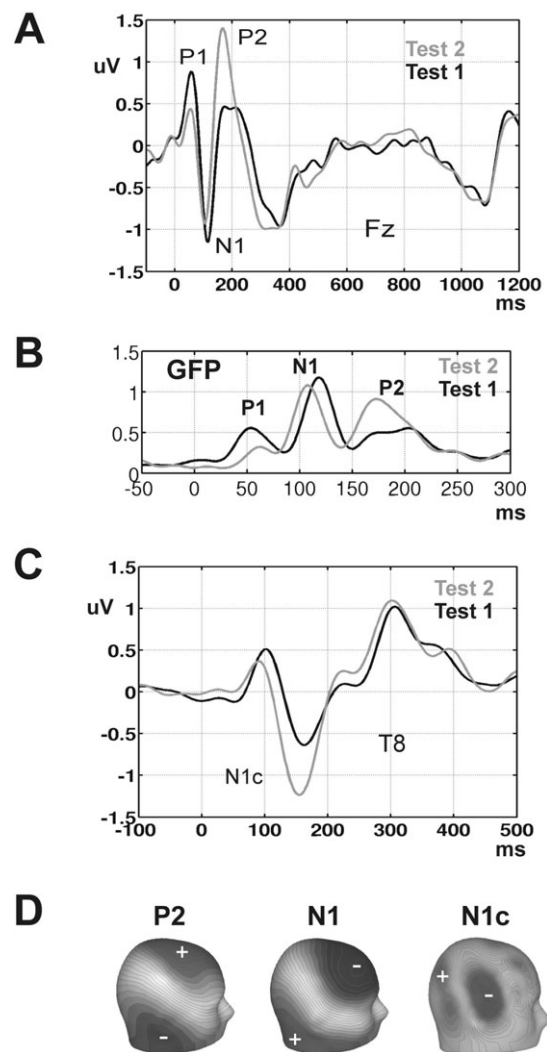


Figure 3. Effect of discrimination training on transient AEPs. (A) Augmentation of the P2 evoked by the trained S1 stimulus. P1, N1 and P2 AEPs are identified and shown for the opening (black, Test 1) and closing (gray, Test 2) test sessions. (B) Global field power evoked by the trained S1 stimulus over all electrodes. (C) Augmentation of the N1c in the right hemisphere by discrimination training. (D) Scalp topographies are given for N1, N1c and P2 at their post-training amplitude maxima.

$P = 0.02$] but not for either of the control sets. When the threshold of discrimination at 2.0 kHz was divided by stimulus frequency after training ($\Delta f/f$), a ratio of 0.46% was found which is similar to ratios reported by discrimination studies using unmodulated tones (He *et al.*, 1998). The slope of the psychophysical function after training was steepest for the 2.0 kHz trained stimulus set and shallowest for the 1.8 kHz control set (Fig. 2C), but differences in slope among the stimulus sets did not reach significance.

Six subjects returned for a retention test on the 2.0 kHz stimulus set 2 months after their last test session. Performance at retention ($P = 0.63$) was lower than on the last training session [$P = 0.75, t(5) = -2.80, P = 0.038$] but remained better than in the first test block [$P = 0.32, t(5) = 3.76, P = 0.013$].

Transient AEPs

N1 and P2 transient responses evoked by the S1 reached their amplitude maxima at frontal electrodes with a polarity reversal at occipital sites. Time domain averages at the frontal electrode

(Fz) and global field power (root mean square of all electrodes) are shown for the trained 2.0 kHz S1 in Figure 3A,B where N1 and P2 components are identified (pre-training latencies of 116 ms and 172 ms, respectively). The early occurring P1 (pre-training latency 57 ms) is also identified in these traces. Scalp topographies are shown for the N1 and P2 at their post-training amplitude maxima in Figure 3D. These results show that discrimination training resulted in an enhancement of P2 amplitude. When referred to the pre-stimulus baseline, P2 amplitude increased from 0.65 μV before training to 1.46 μV after training [$t(7) = 6.03$, $P < 0.001$], corresponding to an increase of 124% for the group as a whole. Enhancement of the P2 was also prominent in global field power (Fig. 3B). On the other hand, N1 and P1 amplitude tended to decrease after training, but these effects did not reach significance.

P2 amplitude is shown before and after training for each stimulus set in Figure 4A, referenced in this case to the peak of the N1 (P2–N1 amplitude) in order to remove influences attributable to variability in the N1. Analysis of variance revealed a main effect of before/after [$F(1,7) = 6.7$, $P = 0.036$] but the interaction of before/after with stimulus set was not significant. When the stimulus sets were examined separately, before/after differences in P2 amplitude were found to be significant only for the trained 2 kHz stimulus set [$t(7) = 4.26$, $P = 0.008$]. However, differences for the control sets were in the direction of training and suggested partial generalization of P2 enhancement to the untrained stimuli. Correlations were calculated between before/after differences in P2–N1 amplitude and the behavioral measure **P** for the trained stimulus set alone, and when the three stimulus sets were combined. These correlations were positive but none reached significance.

Acquisition of the enhanced P2 (referenced to the pre-stimulus baseline) over sessions is shown in Figure 4C which includes training sessions 3 and 13 as well as the opening and closing test sessions. A main effect of sessions was found for this measure [$F(3,21) = 4.15$, $P = 0.019$] which was attributable to increases in P2 amplitude occurring on the 13th session of training and on the closing test session compared to session 3 and pre-training performance ($P < 0.015$ or better). For purposes of comparison, Figure 4C also depicts changes

observed in the amplitude of P1 and N1 responses referenced to their pre-stimulus baselines. Main effects of sessions did not reach significance for either measure ($P = 0.16$ and 0.084 for P1 and N1, respectively).

Figure 3C depicts changes occurring over sessions in a fourth AEP component that reached its amplitude maximum at electrode T8 over the right hemisphere. We identified this surface-negative component as the N1c in accordance with properties described by Woods (1995). The N1c was distinguishable from the N1 and P2 by its radial orientation, by its latency (155 ms) falling between that of these two AEPs, and by its preferential expression in the right hemisphere. Discrimination training enhanced the N1c between the two test sessions for the trained S1 stimulus [$t(7) = 3.81$, $P = 0.007$], gradually over the training series [see Fig. 4C; main effect of sessions $F(3,21) = 4.05$, $P = 0.02$]. Before/after training differences in N1c amplitude for the trained and control stimulus sets are shown in Figure 4B. Although before/after differences were largest for the trained stimulus set, enhancement generalized as well to the 2.2 kHz control set where before/after differences reached significance [$t(7) = 2.89$, $P = 0.023$]. Analysis of variance revealed a main effect of before/after [$F(1,7) = 9.52$, $P = 0.018$], but main effects or interactions involving stimulus set did not reach significance. We also searched for an N1c occurring in the left hemisphere (electrode T7) in each test session. An enhanced polarity-inverted response was observed after training at a peak latency (155 ms) that corresponded with the amplitude maximum of the N1c recorded in the right hemisphere. However, the before/after training difference in the polarity inverted response was not significant at its amplitude maximum ($t = -0.84$), nor were before/after differences detected at any other time point in the T7 trace of the left hemisphere.

We also examined the effect of discrimination training on the latency of the P1, N1, N1c and P2 responses evoked by the trained S1. N1 latency decreased from 116 ms in the first test session to 107 ms in the closing test session, $t(7) = 7.94$, $P < 0.001$. This effect was obtained in every subject and can be seen in Figure 3A,B (time domain traces and global field power). P1 and P2 latency, and N1c latency in the right hemisphere, did not change with discrimination training when measured at their amplitude maxima. However, the leading edge of the P2 and N1c waveforms tended to commence earlier after training compared to their pre-training baselines (see Fig. 3A,C).

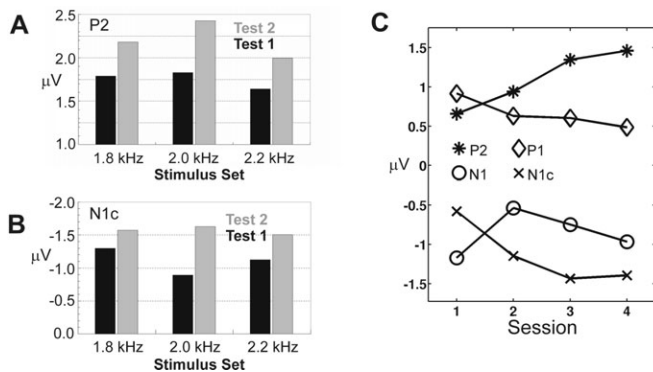


Figure 4. Acquisition of changes in transient AEPs. (A) P2–N1 amplitude is shown on test sessions before and after training for each stimulus set. The 2.0 kHz set was trained. (B) N1c amplitude on test sessions before and after training for each stimulus set. (C) Acquisition of P2 and N1c enhancements over discrimination training. Changes observed for the P1 and N1 are also shown. Response amplitude is referenced to the pre-stimulus baseline and depicted for test session 1, training sessions 3 and 13, and test session 2 (labeled Sessions 1–4, respectively).

Steady State Response

A time domain trace of the 40 Hz SSR evoked by the 2.0 kHz S1 after training is depicted in the lower trace of Figure 1B at its amplitude maximum (Fz electrode). Neither responding at this electrode nor SSR global field power differed between test sessions administered before and after training when calculated over the 1 s S1 period. However, fine grained dynamics were revealed by T^2 when 100 ms windows were moved across the 40 Hz waveform at Fz in 10 ms time steps. Figure 5A gives the results for a representative subject. Two polar plots are shown (right side), each containing vectors depicting SSR amplitude and phase on the 288 accepted test trials in a single 100 ms window before (upper, test 1) and after (lower, test 2) discrimination training. Although phase covers 360° and is variable across single trials, the end point of the mean vector (resultant, shown as the red arrow) is shifted from the origin in

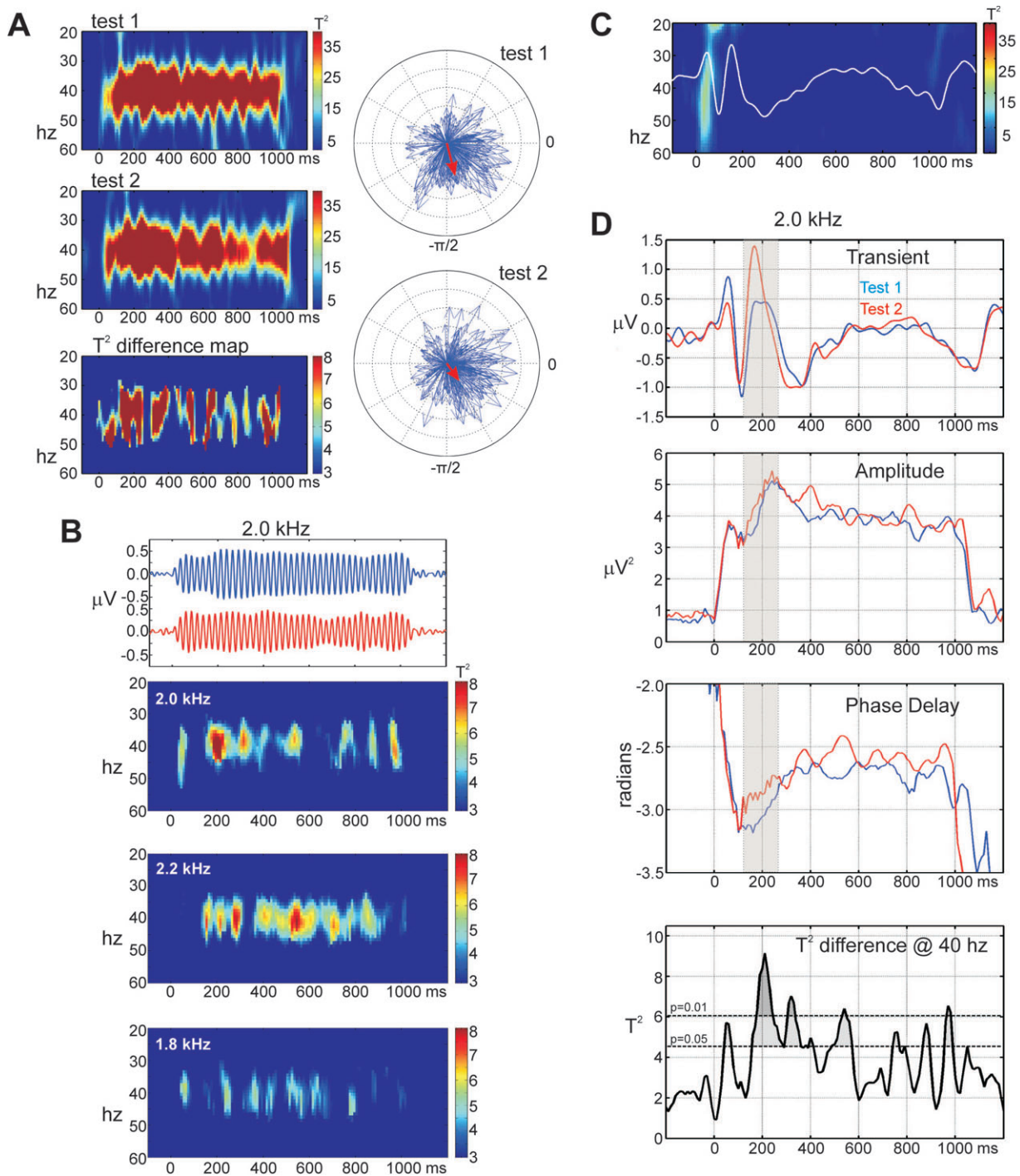


Figure 5. Effect of discrimination training on the 40 Hz SSR. (A) T^2 analysis applied to a representative subject. Polar plots in the right panels depict the 40 Hz SSR in a single 100 ms window 50 ms after onset of the trained S1, on test blocks given before (upper, test 1) and after (lower, test 2) discrimination training. Each vector is a single trial. The spray of vectors is concentrated in the lower right quadrants indicating that an SSR is present (the red arrows are mean vectors). Corresponding T^2 plots are shown in the left panels of the figure (upper panel test 1, middle panel test 2). The lower left panel shows the T^2 difference map obtained for this subject, thresholded for significant before/after differences ($T^2 = 8.0$, $P < 0.01$, red) in the SSR. T^2 differences were more prominent in the first half of the stimulation period. (B) T^2 difference maps are shown for the group as a whole for the trained S1 (2.0 kHz) and control (1.8 and 2.2 kHz) S1 stimuli. Maps are thresholded at $T^2 = 4.5$ ($P < 0.05$, light blue; $T^2 = 6.0$, $P < 0.01$, yellow and above). Time domain traces of the 40 Hz SSR are superimposed above the 2.0 kHz map (dark blue before training, red after). Before/after T^2 differences were observed for the 2.0 kHz S1 in the interval 100–225 ms after stimulus onset with brief periods of significance reappearing thereafter. Generalization to the untrained 2.2 kHz S1 but not to the 1.8 kHz S1 is seen. (C) Evaluation of 40 Hz activity in control subjects tested for discrimination using unmodulated stimuli. The N1/P2 transient response evoked by the unmodulated 2.0 kHz S1 is superimposed on the T^2 map. No 40 Hz activity was observed to overlap the N1/P2 transient waveform evoked by this stimulus. (D) Middle panels: changes in SSR amplitude (mean vector) and phase delay (difference between stimulus and response phase) during the S1 interval on test trials before (test 1) and after (test 2) training, for the group as a whole. For convenience the transient N1/P2 waveform evoked by the trained S1 (top panel) and the group T^2 difference at each time point (bottom panel) are aligned to these data. The shaded areas in the T^2 measure denote level of significance and are extended to the upper panels to identify the region of the largest T^2 effect. A phase advance occurring in the interval 100–225 ms after training was the principal source of the T^2 difference in this interval and in subsequent intervals during the S1 (see phase delay panel). Periods of significance in the T^2 difference map over the duration of the S1 correlated with changes in SSR phase delay ($P < 0.0001$) but not with changes in SSR amplitude measured as the mean vector ($P > 0.40$).

both polar plots, indicating that a 40 Hz SSR is present. Spectral plots of T^2 are shown to the left of Figure 5A and indicate that a 40 Hz SSR was present throughout the stimulation period before (upper plot) and after (middle plot) training (all $T^2 > 45$). The lower spectral plot in Figure 5A shows the T^2 difference between the two test sessions before and after training for this subject, scaled for Monte Carlo significance at $T^2 = 8.0$, $P < 0.01$. Before/after differences reached significance particularly in the first half of the S1 stimulation period, with patches of significance appearing subsequently.

Similar findings were obtained for all subjects to which this analysis was applied. The results are collapsed across subjects in Figure 5B where the before/after T^2 difference is thresholded for significance at $T^2 = 4.5$ ($P < 0.05$, light blue; $T^2 = 6.0$, $P < 0.01$, yellow and above). Results are shown for the trained S1 (2.0 kHz) as well as for the S1s of the untrained 1.8 kHz and 2.2 kHz stimulus sets. Time-domain traces of the 40 Hz SSR evoked by the 2.0 kHz S1 before and after training are superimposed above the T^2 difference map for the 2.0 kHz stimulus. Significant before/after differences were observed in the SSR evoked by the trained S1, particularly in the time interval 150–225 ms after S1 onset, with brief epochs of significance appearing thereafter. Integration of the T^2 statistic over the time interval 50–400 ms at 40 Hz found that before/after differences were stronger for the trained 2.0 kHz S1 than for the untrained 1.8 kHz S1 [$t(5) = 2.29$, $P = 0.035$, one-tailed test] while differences between the 2.0 kHz and 2.2 kHz S1 stimuli were not significant. These results indicate that generalization occurred from training on the 2.0 kHz set to the 2.2 kHz control set, but not to the 1.8 kHz control set.

Augmentation of the SSR within the interval 150–225 ms raises the question of whether the T^2 results shown in Figure 5B might alternatively be attributed to a 40 Hz spectral component of the transient P2 which was also augmented in the vicinity of this time window. To assess this hypothesis, we evaluated 40 Hz activity in the absence of the SSR when N1 and P2 transient responses were evoked by unmodulated 2.0 kHz tones. The results are shown in Figure 5C where the N1/P2 waveform evoked by the unmodulated tone is superimposed on 40 Hz activity evaluated by T^2 at the same scaling used for the upper two T^2 maps of Figure 5A. 40 Hz activity was detected between 30 and 50 ms where middle latency responses or transient gamma band responses were expected (Pantev *et al.*, 1991). However, this activity subsided by ~80 ms and did not extend into the latency window encompassing N1 and P2 transient responses. These findings indicate that T^2 differences observed for the 2.0 kHz AM S1 (Fig. 5B) are not likely to be attributable to a high-frequency component of the enhanced P2 transient response, because no such component was detected in the latency window of the P2 in the unmodulated control condition. Rather, the two responses appeared to be separate brain events.

Changes in the SSR induced by training and detected by T^2 could be generated by changes in the amplitude or phase of the SSR, or both. In order to address this question, we first calculated mean SSR amplitude and phase delay (difference between stimulus phase and response phase) for each subject and 100 ms window during the S1 stimulus. The results are shown in Figure 5D for the group as a whole where SSR amplitude and phase delay (middle panels) are aligned to the transient N1/P2 waveform obtained before and after training (top panel). For convenience, T^2 values comparing group before/after SSR

differences for the trained S1 are plotted over time in the bottom panel of Figure 5D, with light and dark shading indicating $P < 0.05$ and $P < 0.01$, respectively. Light shading is extended into the upper panels of Figure 5 to identify the region of maximum T^2 difference. Inspection of phase delay during the first test session (blue trace) shows that SSR phase shortened gradually commencing ~100 ms post-stimulus and reaching asymptote ~400 ms. After training (red trace) SSR phase advanced by ~0.3 radians (4.8% of the wave period of the SSR) with respect to pre-training performance within this time interval, commencing near but persisting beyond the leading edge of the P2 waveform. Phase advances tended to recur subsequently during the S1 interval, coinciding with significant differences in the T^2 difference map. On the other hand, before/after differences in SSR amplitude were less apparent during the S1 (Fig. 5D, second panel), although a small enhancement is seen during the interval 100–200 ms after stimulus onset. Supplementary analyses not presented in the figure showed that this enhancement was closely paralleled by an increase in phase coherence with no change in absolute vector length, suggesting that it was secondary to a reduction of phase variability around its central tendency during this interval. Multiple regression applied to T^2 differences recorded for the group during the S1 yielded $R = 0.455$ [$F(2,96) = 15.5$, $P < 0.00001$] to which before/after differences in phase contributed [$t(96) = 5.00$, $P < 0.00001$] but differences in mean vector length did not [$t(96) = -0.54$, $P = 0.41$]. These findings indicate that discrimination training modified the temporal properties of the 40 Hz SSR but had little effect on the absolute amplitude of this response. A computer animation showing phase and amplitude dynamics of the mean vector for a representative subject throughout the S1 can be viewed at www.psychology.mcmaster.ca/hnplab.

Source Analyses

The spatial coordinates of regional sources modeled from the grand averaged data for each AEP (N1, N1c, P2 and SSR) were evaluated by analyses of variance collapsing first over before/after test sessions (to examine effects of stimulus set) and then over stimulus sets (to examine effects of before/after). No effects of stimulus set or before/after were found, except for the sources of the P2 which shifted to be more inferior when training had been completed [z coordinate, $F(3,12) = 13.53$, $P = 0.0007$]. However, main effects attributable to AEP were found in both of these analyses. When the six localizations determined for each AEP (three stimulus sets before and after training) were collapsed into a single data set, main effects of AEP were significant for the medial lateral (x) coordinate [$F(3,15) = 25.97$, $P < 0.00001$], anterior-posterior (y) coordinate [$F(3,15) = 9.22$, $P = 0.001$], and inferior-superior (z) coordinate [$F(3,15) = 24.95$, $P < 0.00001$]. The modeled sources for each AEP are co-registered on the average brain of BESA 2000 in Figure 6 in order to visualize their relative positions. *Post hoc* contrasts showed that cortical sources underlying the N1 and N1c were centered lateral with respect to those of the P2 in the region of the auditory cortex ($P < 0.01$ or better, axial view), while sources of the SSR were medial with respect to P2, N1 and N1c sources ($P < 0.03$ or better). P2 sources were also centered anterior with respect to sources of the N1, N1c and SSR ($P < 0.05$ or better), and superior with respect to these sources (minimum $P < 0.0001$) when averaged

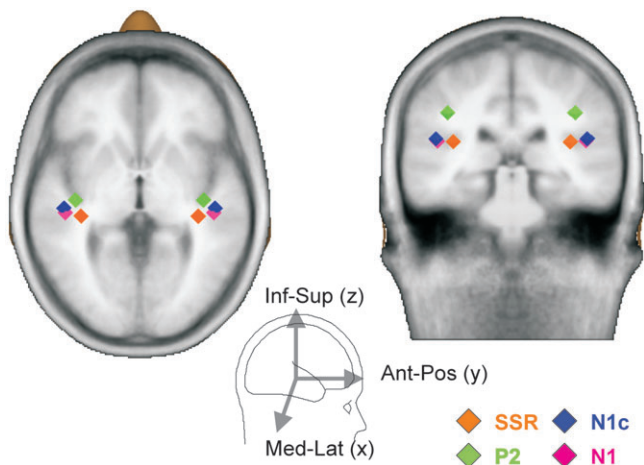


Figure 6. Localizations of the cortical sources of N1, N1c and P2 transient responses and the 40 Hz SSR determined from the grand averaged data. The coordinate system is shown in the inset.

before and after training. These results which confirm SSR sources medial to those of the N1 and P2 are consistent with previous studies which have localized SSR generators by source modeling (Scherg and von Cramon, 1986; Pantev *et al.*, 1996a; Gutschalk *et al.*, 1999; Engelen *et al.*, 2000; Godey *et al.*, 2001; Yvert *et al.*, 2001; Schneider *et al.*, 2002; Shahin *et al.*, 2003) and by intracortical measurements (Celesia, 1976; Liégeois-Chauvel *et al.*, 1993; Godey *et al.*, 2001) to the region of Heschl's gyrus. Differentiation of P2 from N1 and N1c sources and from those of the SSR is in agreement with previous findings which have localized P2 and N1 sources to the region of A2 (Scherg and von Cramon, 1986; Pantev *et al.*, 1996b; Picton *et al.*, 1999) including for P2 sites anterior to the auditory core (Hari *et al.*, 1987; Joutsiniemi *et al.*, 1989; Pantev *et al.*, 1996b). P2 sources may reflect activation centered in anterior auditory belt regions of A2 which receive reciprocal connections from other belt areas and from parabelt zones that project reciprocally to prefrontal cortex (Kaas and Hackett, 1998; Hackett *et al.*, 2001). N1 and N1c sources may reflect activation of posterior and lateral parabelt regions which have dense connections with caudal and rostral parts of the superior temporal gyrus. A note of caution regarding differentiation of P2, N1, and N1c sources within A2 is that source analysis estimates only centers of activation and cannot resolve overlapping generators of similar orientation or determine their spatial extent.

Dipole moment was also contrasted for each AEP before and after training, using the three stimulus sets as the unit of observation. This analysis revealed a main effect of before/after [$F(1,4) = 12.83, P = 0.023$] and an interaction of before/after with AEP [$F(3,12) = 6.331, P = 0.008$]. Both of these effects were attributable to enhanced dipole moments occurring for the P2 in each stimulus set after training [$F(1,4) = 18.06, P = 0.013$] compared with the other AEPs. Dipole moment was not significantly enhanced after training for any other component in either hemisphere. However, subsequent analyses showed that the regional source fitted to the N1 field pattern contained a radially oriented vector that was augmented after discrimination training only in the right hemisphere, with an amplitude peak near 148 ms when the N1 source model was applied to

the N1c time interval. This suggests that dipole moment calculated for the regional source fitted to the N1c field pattern contained contributions arising from the temporally overlapping N1 that obscured changes in radially oriented N1c activity. We also examined the contribution of the three orthogonal vectors of the SSR regional source to the SSR waveform after discrimination training, following the procedure of Scherg and von Cramon (1986). The regional model accounted for 97.9% of the observed field pattern when the three vectors were included. Goodness of fit decreased to 93.3% when only a single tangential source was used to model the field pattern, whereas a single radial source accounted for only 6.3% of the variability in the recorded field pattern. These findings indicate that activity modeled by the tangential vector was the principal contributor to the SSR waveform.

Discussion

We trained non-musician subjects to discriminate small increases in the pitch of a 2.0 kHz standard stimulus, using 40 Hz AM modulated pure tones as the discriminative stimuli. Amplitude modulation allowed us to separate the 40 Hz auditory SSR whose generators localize to the region of Heschl's gyrus in A1 from transient responses of the AEP (N1, N1c, P2) whose modeled centers of activation are spatially differentiable in A2. Discrimination improvement was accompanied by enhancement of the P2 (latency 172 ms) and of the N1c (in the right hemisphere, latency 155 ms), indicating an increase in synchronous neural activity in A2 after training on the discrimination task. The 40 Hz SSR, on the other hand, gave a different picture of cortical dynamics. Overall, there was no overall amplitude enhancement of the SSR; instead we observed a shortening of phase within a latency window coinciding with the onset of the P2 with brief advances in phase reappearing subsequently during the S1. These findings suggest that training at pitch discrimination did not expand the cortical representation for the 2.0 kHz S1 in A1. Instead, temporal properties of the SSR representation were modified by experience on the task. Because both transient and steady-state responses were affected by training, it appears that neural activity was modified in distributed regions of the auditory cortex, particularly in A2 where plasticity appears to be widely expressed in animal studies (Diamond and Weinberger, 1984).

Enhancement of the P2 transient response by acoustic training appears to be a robust phenomenon. To our knowledge, this effect was first described by Eaton and Roberts (1999) in a preliminary study using the present methods. Working independently, Tremblay *et al.* (2001) observed enhancement of the P2 when non-musician subjects were trained to discriminate temporal features of speech signals. More recently, Atienza *et al.* (2002) found an enhancement of the P2 when subjects were trained to detect pitch deviants in a short stream of pitch stimuli. In each of these studies P2 amplitude increased by ~100% when measured from the amplitude peak of the N1 which did not change with training in any study. These results indicate that the neural mechanisms underlying the P2 brain event are sensitive to remodeling by experience. Heretofore this component of the AEP has received little attention in studies of auditory perception, perhaps because in the absence of a training manipulation the P2 shows more limited dynamics.

Enhancement of the N1c by acoustic training has not previously been reported. The expression of the N1c in the right hemisphere in our study where subjects were processing pitch cues is consistent with functional and anatomical evidence for specialization of auditory neurons in this hemisphere for processing of spectral information. Compared to homologous auditory neurons in the left hemisphere, neurons in the right hemisphere are characterized by higher synaptic densities, more closely spaced cortical columns, and comparatively less myelination, which are features that may favor spectral integration of acoustic signals (Zatorre and Belin, 2001). Woods (1995) noted that because its expression is variable, less is known about the N1c component of the AEP compared to other components. A key to expression and enhancement of the right-sided N1c may be the presence of multiple auditory objects in a stimulus sequence that must be distinguished by their spectral properties in order for the subject to comply with task requirements.

In contrast to the P2 and N1c, the N1 (latency 107 ms) was not amplified by discrimination training in our study or in the aforementioned EEG studies of acoustic discrimination. However, enhancement of its magnetic counterpart the N1m by training at pitch discrimination has been reported by Menning *et al.* (2000). It should be noted that an augmented P2 brain event commencing within the N1 latency window would subtract from N1 amplitude in electrical recordings, but not necessarily in magnetic recordings owing to the insensitivity of magnetic sensors to radial currents contributing to the P2. This factor could explain discrepant EEG and MEG findings with regard to N1 amplitude enhancement. Although N1 amplitude was not modified, N1 latency diminished by 9 ms after training in our study. Competition among synapses favoring fast inputs could generate a latency shift of this magnitude (Song *et al.*, 2000), as could an overlapping of AEP components. In the latter respect it may be noteworthy that N1c and P2 responses tended to commence earlier after training within a time interval coinciding with the onset of the N1 (see Fig. 3A,C). When we modeled the N1 field pattern with a regional source, a radial component appeared in the right hemisphere with an early onset latency that could have reflected a contribution arising from the N1c.

The cortical sources that we modeled for the P2, N1, and N1c were consistent with previous studies that differentiated these sources localizing within A2 from sources of the 40 Hz SSR which localize more medially to Heschl's gyrus in the auditory core (Pantev *et al.*, 1993; Schneider *et al.*, 2002; Shahin *et al.*, 2003). However, the changes that we observed in the SSR after training did not include amplitude enhancements that were expected on the basis of research in owl monkeys where increased spiking of A1 neurons (Blake *et al.*, 2002) and expansion of the tonotopic representation in A1 (Recanzone *et al.*, 1993) were found for stimuli associated with reward. Rather, our results are more in line with those of Kilgard *et al.* (2001) which show that behavioral conditioning with multiple frequencies tends to preserve segregated tonotopic representations in A1. Several factors may account for the different findings among these studies including the training procedures that were used, their duration, whether the relevant rules for cortical reorganization were optimized, and whether the methods used to measure cortical reorganization were sensitive to the changes that occurred. With respect to the latter variable it should be noted that our results do not appear to be

attributable to insensitivity of the SSR to the anatomy or functional organization of Heschl's gyrus. Schneider *et al.* (2002) found that the N19-P30 source waveform underlying the SSR was augmented by 102% in musician compared to non-musician subjects, when extracted by deconvolution from AM rates near 39 Hz. The SSR source waveform also correlated highly ($r = 0.87$) with the volume of gray matter in the anteromedial portion of Heschl's gyrus well as with musical aptitude ($r = 0.71$). In our study temporal modulation of the SSR generalized more to the untrained 2.2 kHz S1 than to the untrained 1.8 kHz S1, perhaps because subjects were trained to detect only increases from 2.0 kHz (range 2.0–2.1 kHz) and experienced no stimuli below 2.0 kHz during training. Although behavioral performance did not differ significantly between the two control sets, behavioral performance was consistently better on the 2.2 kHz set as assessed by P , d' , discrimination thresholds, and the slope of psychophysical functions obtained after discrimination training.

Modification of distributed auditory cortical representations in the present study raises the question of how remodeling was achieved and expressed in the AEP. Detailed laminar analyses of multiple unit activity in relation to current sinks and sources in the auditory cortex of the awake monkey suggest that positive-going surface potentials of the P1–N1–P2 complex are generated principally by depolarization of pyramidal neurons in neocortical layers III–VI, while surface negativities reflect depolarization of apical dendrites in the upper neocortical laminae [see Fig. 1 (Fishman *et al.*, 2000) and Fig. 2 (Fishman *et al.*, 1998)]. Results summarized by Mitzdorf (1994) for the cat and for auditory middle latency responses of the rat by Sukov and Barth (1998) are consistent with this interpretation, although a role for hyperpolarization in primate cortex cannot be ruled out (Schroeder *et al.*, 1995). If this interpretation is provisionally accepted for the P2 and N1c components of the human AEP, our results imply that more pyramidal neurons were depolarizing synchronously in A2 after training on the discrimination task than before training commenced. Modulation of the neocortical mantle by the basal forebrain (nucleus basalis magnocellularis, NBM) is one possible source of these enhancements. This structure, which has been implicated in neuroplastic remodeling by many researchers (e.g. Weinberger *et al.*, 1990; Dykes, 1997; Wenk, 1997; Edeline, 1999), contains large cholinergic and GABAergic neurons that project to targets in the neocortex in a broadly tuned corticotopic arrangement (Jiménez-Capdeville *et al.*, 1997). Because GABAergic fibers synapse on inhibitory interneurons (Freund and Meskenaite, 1992), coactivation of cholinergic and GABAergic pathways acts synergistically to increase the sensitivity of pyramidal cells to their afferent inputs, shortening response latency by a magnitude similar to that which we observed in SSR phase after training (Metherate and Ashe, 1993) and strengthening synaptic connections on auditory neurons by a Hebbian correlation rule (Metherate and Weinberger, 1990; Cruikshank and Weinberger, 1996; Kilgard and Merzenich, 1998). These findings suggest that modulation of the neocortical mantle by the NBM serves an attention-like function that gates plastic changes at the synapse and facilitates their expression in performance after synaptic remodeling has occurred. When measured by slow cortical potentials (Pirch, 1993; Pirch *et al.*, 1983), modulation by the NBM has an onset latency resembling that of the auditory N1/P2 complex, as do top-down signals from prefrontal cortex which may

converge on auditory neurons and serve an additional teaching role (Tomita *et al.*, 1999). Although strengthening of modulation itself by conditioning (Rigdon and Pirch, 1986) could account for augmented P2 responses, evidence summarized by Dykes (1997) indicates that additional cortical neurons are likely to become tuned to the task stimuli during training and to contribute to progressive improvements in behavioral performance such as those observed in our study. Network behavior of this nature would be expected to influence plastic remodeling of sensory modalities in addition to audition, although not necessarily at the same latencies observed in the auditory case.

Possible constraints on interpretation of present findings should be acknowledged. Enhancements in the amplitude of P2 and N1c responses could in principle be attributed either to an increase in the number of neurons activated by a stimulus or to an increase in the synchrony of their depolarization. Calculations reported by Hari (1990) suggest that increases in synchronous activity representing 5% of the neurons in a cortical area 1 mm² can account for a scalp recorded AEP. We cannot unequivocally assess the relative contributions of number of neurons or synchrony to enhancement of P2 and N1c transient responses in our study. However, because the temporal envelopes of the augmented P2 and N1c responses were broad and did not appear to change notably after training, an increased number of contributing neurons may have been the more significant variable. Auditory neurons are also sensitive to eye position and the spatial location of acoustic stimuli (Werner-Reiss *et al.*, 2003). This raises the question of whether eye position or head movements induced by the processing of visual feedback cues may have contributed to training effects on AEPs. This would appear to be unlikely, because test sessions before and after training were carried out under identical conditions in which visual feedback cues were eliminated. It is also not clear how undetected head or eye movements directed toward a darkened feedback light in the center of the visual field could preferentially influence the right-sided N1c, or explain P2 enhancements reported in studies by Atienza *et al.* (2002) and Tremblay *et al.* (2001) which used different feedback arrangements (feedback after only blocks of trials, or no feedback during testing, respectively).

Shahin *et al.* (2003) recently reported that P2 responses evoked by musical tones in violinists and pianists were larger than those observed in non-musician subjects, as were right-sided N1cs. These results could have been predicted from the present findings owing to the different training histories of musicians and non-musicians with respect to tones of musical timbre. On the other hand, our findings with regard to the effects of training on the 40 Hz SSR suggest a dissociation of transient and SSR components of the AEP, with neuroplastic transient responses expressing as amplitude enhancements in training studies and in musicians but 40 Hz SSR enhancements in musicians only (Schneider *et al.*, 2002) where they may be an anatomical marker for musical skill. However, we cannot exclude the possibility that other training procedures may modify SSR amplitude and its anatomical substrate, depending on the type of training that is given, its duration, and when it is delivered in the course of brain development.

Notes

This research was supported by grants from the Canadian Institutes of Health Research (Operating and NET) and the Natural Sciences and Engineering Research Council of Canada.

Address correspondence to L.E. Roberts, Department of Psychology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. Email: roberts@mcmaster.ca.

References

- Atienza M, Cantero JL, Dominguez-Marin E (2002) The time course of neural changes underlying auditory perceptual learning. *Learn Mem* 9:138–150.
- Beitel RE, Schreiner CE, Cheung SW, Wang X, Merzenich MM (2003) Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. *Proc Natl Acad Sci USA* 100:11070–11075.
- Blake DT, Strata F, Churchland AK, Merzenich MM (2002) Neural correlates of instrumental learning in primary auditory cortex. *Proc Natl Acad Sci USA*. 99:10114–10119.
- Buonomano DV, Merzenich MM (1998) Cortical plasticity: from synapses to maps. *Annu Rev Neurosci* 21:149–186.
- Celesia GG (1976) Organization of auditory cortical areas in man. *Brain* 99:403–414.
- Cruikshank SJ, Weinberger NM (1996) Receptive-field plasticity in the adult auditory cortex induced by Hebbian Covariance. *J Neurosci* 16:861–875.
- Dember WN, Warm JS (1979) *Psychology of perception*. New York: Holt Rinehart Winston.
- Diamond DM, Weinberger NM (1984) Physiological plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response: II. Secondary field (AII). *Behav Neurosci* 98:189–210.
- Dykes RW (1997) Mechanisms controlling neuronal plasticity in somatosensory cortex. *Can J Physiol (Lond) Pharmacol* 75:535–545.
- Eaton RA, Roberts LE (1999) Effect of spectral frequency discrimination on auditory transient and steady state responses in humans. *Soc Neurosci Abstr* 29:156.15.
- Edeline J (1999) Learning-induced physiological plasticity in the thalamo-cortical sensory systems: a critical evaluation of receptive field plasticity, map changes and their potential mechanisms. *Prog Neurobiol* 57:165–224.
- Engelien A, Schulz M, Ross B, Arolt V, Pantev C (2000) A combined functional *in vivo* measure for primary and secondary auditory cortices. *Hear Res* 148:153–160.
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (1998) Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey. *Brain Res* 786:18–30.
- Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2000) Complex tone processing in primary auditory cortex of the awake monkey. II. Pitch versus critical band representation. *J Acoust Soc Am* 108:247–262.
- Freund TF, Meskenaite V (1992) γ -Aminobutyric acid-containing basal forebrain neurons innervate inhibitory interneurons in the neocortex. *Proc Natl Acad Sci USA* 89:738–742.
- Galambos R, Makeig S, Talmachoff P (1981) A 40-Hz auditory potential recorded from human scalp. *Proc Natl Acad Sci USA* 78:2643–2647.
- Godey B, Schwartz D, de Graaf JB, Chauvel P, Liégeois-Chauvel C (2001) Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin Neurophysiol* 112:1850–1859.
- Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. New York: Wiley.
- Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hähnel S, Picton TW, Scherg M (1999) Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clin Neurophysiol* 110:856–868.
- Hackett TA, Preuss TM, Kaas JH (2001) Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 441:197–222.

- Hari R (1990) The neuromagnetic method in the study of the human auditory cortex. In: Auditory evoked magnetic fields and electric potentials. *Advances in Audiology* (Grandori F, Hoke M, Romani GL, eds), vol. 6, pp. 222–282. Basel: Karger.
- Hari R, Pelizzzone M, Makeka JP, Hallstrom J, Leinonen L, Lounasmaa OV (1987) Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology* 26:31–43.
- He N, Dubno JR, Mills JH (1998) Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. *J Acoust Soc Am* 103:553–565.
- Jesteadt W, Bilger RC (1974) Intensity and frequency discrimination in one- and two-interval paradigms. *J Acoust Soc Am* 55:1266–1276.
- Jiménez-Capdeville ME, Dykes RW, Myasnikov AA (1997) Differential control of cortical activity by the basal forebrain in rats: a role for both cholinergic and inhibitory influences. *J Comp Neurol* 381:53–67.
- John MS, Dimitrijevic A, Picton TW (2001) Weighted averaging of steady-state responses. *Clin Neurophysiol* 112:555–562.
- Joutsiniemi SL, Hari R, Vilkmann V (1989) Cerebral magnetic responses to noise bursts and pauses of different durations. *Audiology* 28:325–333.
- Kaas JH, Hackett TA (1998) Subdivisions of auditory cortex and levels of processing in primates. *Audiol Neurootol* 3:73–85.
- Kilgard MP, Merzenich MM (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279:1714–1718.
- Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J Neurophysiol* 86:326–338.
- Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P (1993) Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol* 92:204–214.
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477.
- Manly BFJ (1991) *Randomization and Monte Carlo methods in biology*. New York: Chapman & Hall.
- Menning H, Roberts LE, Pantev C (2000) Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport* 11:817–822.
- Metherate R, Ashe JH (1993) Nucleus basalis stimulation facilitates thalamocortical synaptic transmission in the rat auditory cortex. *Synapse* 14:132–143.
- Metherate R, Weinberger NM (1990) Cholinergic modulation of responses to single tones produces tone-specific receptive field alterations in cat auditory cortex. *Synapse* 6:133–145.
- Mitzdorf U (1994) Properties of cortical generators of event-related potentials. *Pharmacopsychiatry* 27:49–52.
- Pantev C, Makeig S, Hoke M, Galambos R, Hampson S, Gallen C (1991) Human auditory evoked gamma-band magnetic fields. *Proc Natl Acad Sci USA* 88:8996–9000.
- Pantev C, Elbert T, Makeig S, Hampson S, Eulitz C, Hoke M (1993) Relationship of transient and steady-state auditory evoked fields. *Electroencephalogr Clin Neurophysiol* 88:389–396.
- Pantev C, Roberts LE, Elbert T, Ross B, Wienbruch C (1996a) Tonotopic organization of the sources of human auditory steady-state responses. *Hear Res* 101:62–74.
- Pantev C, Eulitz C, Hampson S, Ross B, Roberts LE (1996b) The auditory evoked ‘off’ response: sources and comparison with the ‘on’ and the ‘sustained’ responses. *Ear Hear* 17:255–265.
- Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M (1998) Increased auditory cortical representation in musicians. *Nature* 392:811–814.
- Picton TW, Vajsar J, Rodriguez R, Campbell KB (1987) Reliability estimates for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol* 68:119–131.
- Picton TW, Alain A, Woods DL, John MS, Scherg M, Valdes-Sosa P, Bosch-Bayard J, Trujillo NJ (1999) Intracerebral sources of human auditory-evoked potentials. *Audiol Neurootol* 4:64–79.
- Picton TW, Dimitrijevic A, John MS, Van Roon P (2001) The use of phase in the detection of auditory steady-state responses. *Clin Neurophysiol* 112:1698–1711.
- Pirch JH (1993) Basal forebrain and frontal cortex neuron responses during visual discrimination in the rat. *Brain Res Bull* 31:73–83.
- Pirch JH, Corbus MJ, Rigdon GC (1983) Single-unit and slow potential responses from rat frontal cortex during associative conditioning. *Exp Neurol* 82:118–130.
- 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.
- Rigdon GC, Pirch JH (1986) Nucleus basalis involvement in conditioned neuronal responses in the rat frontal cortex. *J Neurosci* 6:2535–2542.
- Scherg M, von Cramon D (1986) Evoked dipole source potentials of the human auditory cortex. *Electroencephalogr Clin Neurophysiol* 65:344–360.
- Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A (2002) Morphology of Heschl’s gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci* 5:688–694.
- Schroeder CE, Steinschneider M, Javitt DC, Tenke CE, Givre SJ, Mehta AD, Simpson GV, Arezzo JC, Vaughan HG (1995) Localization of ERP generators and identification of underlying neural processes. *Electroencephalogr Clin Neurophysiol* 44(Suppl.):55–75.
- Shahin A, Bosnyak DJ, Trainor, DJ, and Roberts LE (2003) Enhancement of Neuroplastic P2 and N1c Auditory evoked potentials in musicians. *J Neurosci* 23:5545–5552.
- Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* 3:919–926.
- Sukov W, Barth DS (1998) Three-dimensional analysis of spontaneous and thalamically evoked gamma oscillations in auditory cortex. *J Neurophysiol* 79:2875–2884.
- Timm NH (1975) *Multivariate analysis with applications in education and psychology*. Monterey, California: Brooks/Cole.
- Tomita H, Ohbayashi M, Nakahara, K, Hasegawa I, Miyashita, Y (1999) Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401:699–703.
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B (2001) Central auditory plasticity: changes in the N1–P2 complex after speech-sound training. *Ear Hear* 22:79–90.
- Valdes-Sosa MJ, Bobes MA, Perez-Abalo MC, Perera M, Carballo JA, Valdes-Sosa P (1987) Comparison of auditory-evoked potential detection methods using signal detection theory. *Audiology* 26:166–178.
- Victor JD, Mast J (1991) A new statistic for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol* 78:378–388.
- Wenk GL (1997) The nucleus basalis magnocellularis cholinergic system: one hundred years of progress. *Neurobiol Learn Mem* 67:85–95.
- Weinberger NM, Ashe JH, Metherate R, McKenna T.M., Diamond D.M., Bakin JS, Lennartz RC, Cassady JM (1990) Neural adaptive information processing: a preliminary model of receptive-field plasticity in auditory cortex during Pavlovian conditioning. In: *Learning and computational neuroscience: foundations of adaptive networks* (Gabriel M, Moore J, eds), pp. 91–138. Cambridge, MA: MIT Press.
- Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM (2003) Eye position affects activity in primary auditory cortex of primates. *Curr Biol* 13:554–562.
- Woods DL (1995) The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalogr Clin Neurophysiol* 44(Suppl.):102–109.
- Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C (2001) Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb Cortex* 11:411–423.
- Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11:946–953.