Published in final edited form as:

Clin Neurophysiol. 2009 January ; 120(1): 128–135. doi:10.1016/j.clinph.2008.10.005.

Auditory training alters the physiological detection of stimulus-specific cues in humans

Kelly L. Tremblay, Ph.D.¹, Antoine J. Shahin, Ph.D.², Terence Picton, M.D., Ph.D.³, and Bernhard Ross, Ph.D.³

¹ University of Washington, Dept. of Speech and Hearing Sciences, Seattle, WA, USA

² University of California Davis, Center for Mind and Brain, Davis, CA, USA

³ Rotman Research Institute, Baycrest Centre and University of Toronto, Toronto, Canada

Abstract

Objective—Auditory training alters neural activity in humans but it is unknown if these alterations are specific to the trained cue. The objective of this study was to determine if enhanced cortical activity was specific to the trained voice-onset-time (VOT) stimuli ‘mba’ and ‘ba’, or whether it generalized to the control stimulus ‘a’ that did not contain the trained cue.

Methods—Thirteen adults were trained to identify a 10 ms VOT cue that differentiated the two experimental stimuli. We recorded event-related potentials (ERPs) evoked by three different speech sounds ‘ba’ ‘mba’ and ‘a’ before and after six days of VOT training.

Results—The P2 wave increased in amplitude after training for both control and experimental stimuli, but the effects differed between stimulus conditions. Whereas the effects of training on P2 amplitude were greatest in the left hemisphere for the trained stimuli, enhanced P2 activity was seen in both hemispheres for the control stimulus. In addition, subjects with enhanced pre-training N1 amplitudes were more responsive to training and showed the most perceptual improvement.

Conclusion—Both stimulus-specific and general effects of training can be measured in humans. An individual’s pre-training N1 response might predict their capacity for improvement.

Significance—N1 and P2 responses can be used to examine physiological correlates of human auditory perceptual learning.

Keywords

Auditory learning; Auditory Plasticity; Auditory training; P2; Speech training

INTRODUCTION

The central auditory system changes as a function of experience, reorganizing throughout the lifespan according to available auditory input. One way of shaping the auditory system is to use auditory training exercises. Animal studies have shown that auditory processing can be altered with training and such changes have been attributed to a number of processes including:

Corresponding Author: Kelly L. Tremblay, Ph.D. Associate Professor, Dept. of Speech and Hearing Sciences University of Washington, Seattle, WA 98115, tremblay@u.washington.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1) greater numbers of neurons responding in the sensory field, 2) improved neural synchrony (or temporal coherence), and 3) de-correlated activity among neurons, whereby each neuron responds differently in relation to their functional specificity relative to other members of the population (Barlow and Foldiak, 1989; Gilbert et al., 2001)

Scalp-recorded brain activity (EEG) has been used to measure training-related changes in humans. In a series of studies, we trained naive listeners to identify two within category pre-voiced 'ba' sounds differing in voice-onset-time (VOT) (Figure 1). Following training, and coinciding with improved perception, the magnitude of the auditory evoked response increased (Tremblay et al., 1997; Tremblay et al., 2001). Although similar experience-related changes in evoked response morphology have been reported (Atienza et al., 2002; Menning et al., 2002; Reinke et al., 2003; Shahin et al., 2003; Bosnyak et al., 2004; Shahin et al., 2005; Alain et al., 2007) little is known about the underlying neural mechanisms contributing to these surface recorded physiological changes, or how they contribute to perception (for review, see Tremblay, 2007)

There is evidence to suggest that training alters the sensory encoding of the specific trained cue(s) and that these physiological changes might in turn contribute to improved perception (for reviews see, Irvine and Wright, 2005; Dahmen and King., 2007; Fritz et al., 2007). Training-related perceptual gains seen in children with learning and language disorders, for example, have been attributed to improved temporal coherence of neurons (neural synchrony) representing the specific cues emphasized during listening training. Auditory training is also used as a rehabilitation tool for people with hearing loss and who use cochlear implants or hearing aids. Even though auditory training paradigms are often used in clinics to improve the perception of certain cues, by using specific stimuli (or modifications thereof) and specific tasks, there is little evidence to support or refute if listening training alters the physiological detection of specific acoustic cues. What's more, there is evidence to suggest that processes such as attention and arousal (Amitay et al., 2006) or mere stimulus exposure (Sheehan et al., 2005) account for some of the training-related changes reported in the literature.

Therefore, to learn more about the functional significance of the physiological changes following training, we asked three questions: 1) Are physiological changes seen in all individuals and do they relate to changes in perception? 2) Are the physiological enhancements specific to the cue being trained? and 3) Can patterns of physiological enhancement tell us something about how the auditory system is affected by training? To answer these questions, multiple control conditions were added to our previous research designs. First, the training program was shortened to provide the opportunity to observe "non-learners". "Learners" and "non-learners" participate in the same training task and experience similar stimuli; therefore, observed physiological changes seen in "non-learners" could represent simple exposure to the stimuli. To determine if the training-related changes are specific to the VOT cue being trained, we examined if post-training physiological enhancements were seen in response to the vowel 'a', the portion of the stimulus that did not contain the trained VOT cue. Patterns of evoked activity were examined because asymmetric changes, following VOT training, have been reported (Tremblay et al. 1997; Tremblay & Kraus 2002), and we questioned if the reported laterality effect would be observed when intra-cerebral source analysis was used. Moreover, would the laterality effects for the trained VOT stimuli differ from that evoked by the control stimulus /a/? Our hypothesis was that auditory training would differentially affect the responses evoked by the experimental and control stimuli.

METHODS

Participants

Thirteen young normal-hearing, mono-lingual English speaking, right-handed adults (age 21–30 years) participated in this experiment. They were in good general health, reported no history of otological or neurological disorders, and provided their written consent (University of Washington approved form) prior to participating in this experiment.

Stimuli

Two Klatt synthesized (Klatt, 1980) pre-voiced ‘ba’ stimuli were used in this experiment (Figure 1). They are the same stimuli used in our previous experiments (Tremblay et al., 1997; 1998; 2001; Tremblay and Kraus, 2002); therefore, additional stimulus descriptions can be found in our previous publications. Spectrally the two stimuli are identical, but they differ in terms of VOT. One stimulus has a VOT of –20 ms while the other is –10 ms. Without training, native English speakers routinely categorize these two pre-voiced stimuli ‘ba’ (McClaskey et al., 1983; Tremblay et al., 1997; 1998; 2001). The purpose of the training task was to teach individuals to detect the VOT difference between the two stimuli and identify the –20 ms VOT stimulus as ‘mba’, and the –10 ms VOT stimulus as ‘ba’.

A vowel stimulus that did not contain the VOT cue was created as a control condition. Using Neuroscan Stim-Sound software (version 3.7), the vowel stimulus was created by segmenting and deleting the consonant portion (45 ms) of the –20 ms VOT stimulus and then windowing the first 10 ms of the onset of the steady state vowel ‘a’ using a Hanning type window.

A brief period of silence precedes the onset of each sound (approximately 50 ms for the –20 ms stimulus; 60 ms for the –10 ms VOT token, and 105 ms for the ‘a’). Because of these silent periods, evoked response peak latencies are delayed by this same amount of time.

Procedure

Each individual participated in eight sessions (Figure 2). A pre-training test session on day one, was followed by six training sessions (days 2–7) and a post-training test session on day eight. Sessions did not always take place on consecutive days. The same stimulation equipment and same intensity settings were used during all sessions.

Test sessions

Electrophysiology testing and analyses: Pre- and post-training test sessions began with electrophysiological testing. Participants did not perform a task during EEG recording; instead, they were asked to ignore the stimuli and watch a silent closed-captioned video of their choice. Each stimulus (e.g. ‘mba’) was presented 500 times in a single block of trials. The procedure was repeated for each stimulus type (‘ba’, ‘mba’ and ‘a’) in randomized order. Blocks of homogeneous stimuli were used to optimize N1 and P2 recordings and minimize overlapping discriminative processes. The stimulus onset asynchrony (SOA) was 1175ms. Stimuli were presented monaurally to the right ear at a level of 74 dB SPL. Evoked potential activity was filtered on-line from 0.15 to 100 Hz (12 dB/octave roll off) and recorded using a 32-channel Neuroscan™ Quik-Cap system. The ground electrode was placed on the forehead and the reference electrode was on the nose. Eye blink activity was monitored using an additional channel with electrodes located superiorly and inferiorly to one eye and at the outer canthi of both eyes. Trials with ocular artifacts exceeding ± 70 microvolts were rejected from averaging. Approximately 20 percent of all trials were rejected and the remaining sweeps were averaged and filtered off-line from 1 Hz (high-pass filter, 24dB/octave) to 20 Hz (low-pass filter, 24 dB/octave).

Because there is interest in developing a time efficient tool for examining the effects of auditory training in clinical settings, where the application of many electrodes might not be feasible, investigators often record and analyze evoked brain activity from a single electrode (e.g., Cz). However, it is also important to know if the interpretations based on one type of analysis (e.g., Cz) are consistent when information from other electrode sites are considered. For this reason, we analyzed pre- and post-training recordings in three ways: 1) from electrode site Cz, 2) global field power measures (Skrandies, 2003), and 3) hemispheric differences using source analysis. Helmholtz tests of contrast (Harville 1997), in which the mean of one variable (e.g., control 'a' stimulus) is compared to another variable with more than one level (e.g., experimental 'mba' and 'ba' conditions) were used to compare stimulus effects, across days (pre- and post-training). When comparing the effects of training using source waveforms, Helmholtz tests of contrasts were used to compare the effects of stimulus type (2 levels: control vs. experimental); and training (2 levels: pre- and post-training) across hemispheres (2 levels: left vs. right).

1) Cz Peak Analysis: To compare with previously published literature; recordings from electrode Cz were examined. In all instances, each peak (e.g., P2) was analyzed separately and peak amplitude was calculated relative to pre-stimulus baseline. For each individual, peak latency and amplitude values were selected by identifying the maximum or minimum peaks within a specified latency region (± 20 ms) that was based on group averaged data from midline electrodes.

2) Global Field Power (GFP) Analysis: GFP measures, which quantify the instantaneous global brain activity across the entire scalp, were used to examine P1, N1, P2 responses. Amplitudes and latencies were based on each participant's GFP waveform, for each stimulus type, and each recording condition (pre- and post-training), using a 20 ms latency window around each peak that was derived from group averaged data.

3) Hemispheric Source Analysis: Source analysis was performed on the grouped ($n = 13$) data for each stimulus ('a', 'ba', 'mba') and training conditions (pre-training, post-training) using BESA software (Brain Electrical Source Analysis, MEGIS Software, Graefelfing - Germany). A four-shell ellipsoidal head model was used for source analysis. Two regional sources (one for each auditory cortex) were used to fit each component (N1 and P2). Mirrored symmetric source locations with respect to the mid-sagittal plane were used; however, similar results were also obtained without symmetry constraints (the right regional source fitted about 1 cm anterior to the left). Including constraints (e.g. anatomical, physiological) as part of the inverse problem in source analysis, reduces the number of possible solutions and enhances the likelihood that a unique solution is achieved (Scherg, 1990). Because the results were similar for both symmetrical and asymmetrical source solutions, we only present the symmetrical in the results. The sources were fit using windows ± 20 ms around the minimum of the N1, or the maximum of the P2 peak. A regional source in EEG models activity arising from fissures as well as gyri and consists of three orthogonal vectors (Scherg, 1990) - radial (medial-lateral), first tangential (inferior-superior), and a second tangential (anterior-posterior). Modeling radial and tangential activity by the regional source allows for "approximation for the whole electric scalp activity arising from a cortical region with a maximal extension of some 2–3 cm" (Scherg, 1990).

Because N1 and P2 source locations for conditions were localized very close to each other (mean coordinates for the N1 and P2 source locations across conditions in Talairach space were: N1($x = 48$ mm, $y = -20$ mm, $z = 17$ mm) and P2($x = 45$ mm, $y = -21$ mm, $z = 17$ mm)) a global source model for the N1-P2 wave was obtained as the average across N1 and P2 source coordinates across stimulus conditions. The source model was applied back as a spatial filter onto each subject data to obtain the corresponding individual source waveforms for each condition and measurement. N1 and P2 peak amplitude and latencies were measured from source waveforms for radial (R), first (T1), and second tangential (T2) components for each

regional source using Matlab. P1, N1 and P2 amplitudes were determined for each participant, each stimulus type, and each recording condition (pre- and post-testing) for each hemisphere using a 40 ms window centered about the group defined peak for each condition.

Behavioral testing and analyses: Baseline measures were obtained using an identification task following the first electrophysiological recording. When each stimulus was presented, the participant was asked to label the sound they heard. Two choices were provided on the computer screen: 'mba' and 'ba'. The response was scored correct if the subject assigned 'mba' to the -20 ms VOT stimulus and 'ba' to the -10 ms VOT stimulus. The task was self paced. The participant did not receive feedback after each trial, but they were told their total score after completing the block of 50 trials (25 tokens of each stimulus). Stimuli were presented binaurally at a level of 74 dB SPL. Performance scores were reported as estimates of d-prime derived from the hits, misses, false alarms, and correct rejection rates (MacMillan and Creelman, 1991). These results served as identification test scores for pre- (Day 1) and post-training (Day 8) test sessions.

Training Sessions —On day 2, subjects were given an easy VOT contrast. They were asked to identify a -30 ms VOT stimulus from a -10 ms VOT stimulus. Our prior studies show this 20 ms pre-voiced distinction is an easy contrast for untrained listeners to perceive; therefore, this session allowed the subjects to listen to the pre-voiced stimuli and orient themselves to the pre-voiced cue using an easier stimulus pair. Feedback, in the form of a green light appeared on the computer screen, when the subject correctly identified the -30 ms VOT stimuli as 'mba' and the -10 ms VOT stimuli as 'ba'. After this initial session, each subject began training using the -20 ms and -10 ms VOT stimuli. On each day, identification training sessions consisted of four blocks of 50 trials in which either a -10 ms or -20 ms VOT stimulus was presented. Once again, feedback (green reinforcement light) was given when the -20 ms VOT stimuli was labeled as 'mba' and the -10 ms VOT labeled as 'ba'. Each stimulus was presented randomly with an equal probability of occurrence. Participants were allowed to view their score at the end of each block of trials.

RESULTS

I. Effects of Training

Perception—The ability to identify the two sounds improved with training ($t = 4.75$, $df = 12$, $p = 0.0005$). Despite the increase in averaged performance scores, there was variability across individuals (Figure 3). Three individuals could be described as non-learners, indicated by the symbol 'n', because their performance declined or did not improve beyond test re-test reliability levels (Tremblay et al., 2001).

Electrophysiology—The effects of training on the EP waveforms are shown in Figures 4 and 5 and the source waveforms are shown in Figure 6. As described in the methods, peak latencies are delayed by the 50–105 ms of silence that preceded each stimulus onset.

Cz results: Group averaged evoked responses for selected electrodes are shown in Figure 4. P2 amplitude increased following training ($F = 19.1$, $df = 1, 12$, $p = 0.001$) and the training x stimulus type interaction approached but did not reach significance ($F = 2.78$, $df = 1, 12$, $p = 0.12$). There were no significant main effects of training for P1 or N1 amplitude.

GFP results: As shown in Figure 5, significant increases in P2 amplitude can be seen for each stimulus type ($F = 7.87$, $df = 1, 12$, $p = 0.02$), but there was no stimulus x training interaction for P2 ($F = 0.32$, $df = 1, 12$, $p = 0.85$) and no significant training effects for P1 or N1.

Source and hemispheric comparisons (Figure 6A): N1: There were no significant main effects or interactions, involving training, regardless of source (T1, R, or T2), except for astimulus x train x hemisphere interaction for T2 ($F = 8.41$, $df=1,12$ $p = 0.01$). P2: A significant effect of training was found for P2 when modeled using the T1 source ($F = 14.57$, $df = 1,12$, $p=0.002$), and these patterns of change appear to differ depending on the stimulus and hemisphere. As can be seen in Figure 6B, there was a significant effect of hemisphere ($F=11.75$, $df= 1,12$ $p = 0.005$), with P2 amplitude being larger over the left hemisphere, contralateral to the ear of stimulation (L-R mean difference = 6.9). There was also a train x hemisphere interaction ($F = 20.17$, $df=1,12$ $p= 0.001$) with a post-pre train mean difference of 7.05 μv over the left hemisphere, and 3.7 μv over the right hemisphere. Most importantly there was a stimulus x hemisphere x train interaction ($F = 5.07$, $df=1,12$ $p = 0.04$). This effect is most evident for right hemisphere recordings; the amount of P2 enhancement for the experimental stimuli (post-pre mean difference = 2.84 μv) is less than the amount of P2 change for the control stimuli (post-pre mean difference = 5.89 μv). When other sources (R or T2) were considered, no significant P2 effects or interactions for training were obtained.

II. Brain Behavior Relationships

Because pre-training N1 amplitudes appeared to be larger for learners than for non-learners at electrode site Cz (Figure 7B), a linear regression, plotting individual N1 pre-training amplitudes as a function of change in performance (defined by the post- minus pre-training d-prime scores), was conducted (Figure 7A). Significant correlations were found for ‘mba’ ($r = 0.6$, $p = 0.02$), ‘ba’ ($r = 0.6$, $p = 0.04$), and ‘a’ type stimuli ($r = 0.45$ $p = 0.10$), suggesting that the strength of an individual person’s N1 when recorded from vertex might predict the capacity for improvement with this particular training program. Put another way, the larger a person’s pre-training N1 amplitude, the greater the perceptual change following training. Similar correlations were obtained when comparing pre-training N1 GFP measures to d-prime change scores for ‘mba’ ($r = 0.55$, $p = 0.05$), but less so for ‘ba’ ($r = 0.28$, $p = 0.36$) or ‘a’ ($r = 0.43$, $p = 0.13$).

Despite the fact that post-training P2 amplitudes are enhanced, and there appears to be differential effects for non-learners and learners, there were no significant correlations between the amount of P2 amplitude change and the amount of perceptual change when GFP and Cz recordings were analyzed. However, an interesting left hemisphere finding was obtained. For all stimulus conditions, people who showed the largest change in P2 amplitude over the left hemisphere were also the people who started off with the poorest pre-training d-prime scores (‘a’ $r = 0.54$, $p = 0.05$; ‘mba’ $r = 0.56$, $p = 0.04$; ‘ba’ $r = 0.65$, $p = 0.03$). These relationships were not seen for right hemisphere P2 source waveforms.

Results Summary—To determine if the training-related changes are specific to the VOT cue being trained, we examined if post-training physiological enhancements were seen in response to the control stimulus ‘a’, the portion of the stimulus that did not contain the trained VOT cue. Significant increases in P2 amplitude were seen for both the experimental and control stimuli when recorded from vertex (Cz), GFP, and source measures. But the distribution of P2 change was different for the control and experimental stimulus conditions. Whereas increases in P2 amplitude were seen across both hemispheres for the control stimulus ‘a’, post-training P2 responses were larger over the left hemisphere for the experimental stimuli. Although changes in perception did not significantly correlate with changes in P2 amplitude in all stimulus conditions, the amount of P2 change over the left hemisphere for the control stimulus did significantly correlate with the amount of perceptual improvement. Another significant finding was that people with smaller N1 amplitudes were less affected by training, showing little or no perceptual gains following training.

DISCUSSION

Auditory training paradigms are often designed to improve the perception of certain cues, by using specific stimuli (or modifications thereof) and specific tasks. As an example, stimuli differing in VOT are used to train the perception of VOT. An assumption is that the perceptual and physiological changes that occur during training are in part specific to the stimuli and the task.

Because animal studies have shown that time-varying acoustic cues such as VOT are faithfully represented in the timing patterns of neurons (Eggermont, 1995; Steinschneider et al., 1995) it follows that focused listening tasks, using stimuli that vary in VOT, might improve the timing codes responsible for conveying this acoustic information. As an example, we reported significant enhancements in the P2 event-related potential following within category VOT training (Tremblay 2001). Because the reported physiological changes coincided with improved perception of the trained stimuli, one interpretation of these findings was that enhanced P2 amplitudes reflect training-related changes in the temporal coherence of neurons (neural synchrony) representing the distinguishing VOT cue. However, it is also possible that the post-training P2 changes reflect other processes that are activated during testing and training, which are not specific to the trained cue. For this reason, in this study, we questioned if post-training enhanced P2 activity was specific to the VOT cue that was trained.

At first glance, when analyzing midline electrodes and GFP measures, the training-related physiological changes reported here do not appear to be stimulus-specific because increases in P2 amplitude were seen in response to the experimental and control stimuli. Because the control stimulus 'a' did not contain the consonant portion of the (VOT) cue, we could conclude that training did not specifically alter the timing relationship between the consonant and vowel. Another explanation would be that the effects of training "generalized" to other stimuli (e.g., 'a') that share common acoustic features, because the 'a' stimulus shares vowel frequencies with the other two stimuli and has a similar short duration. However this interpretation is based on patterns of brain activity recorded from a single midline electrode (the site typically reported in the literature), as well as GFP measures. When patterns of brain activity across hemispheres are taken into consideration, a second story emerges.

Previously we reported increases in P2 amplitude following training when measured from a subset of electrodes over left and right frontal cortices (Tremblay and Kraus, 2002). This same effect can be seen in Figure 4. However, when source waveforms for left and right hemispheres are examined, stimulus specific physiological changes can be seen. This is especially true for the main part of the tangential component (T1). For the experimental stimuli, post-training enhancements in P2 source amplitude were most evident over the left hemisphere. This laterality finding cannot easily be dismissed as reflecting stronger, training-related, activation patterns contralateral to the ear that was stimulated during evoked potential testing because the 'a' stimulus was also delivered monaurally to the right ear and the ERP to the 'a' changed in both hemispheres.

When asymmetrical changes in brain activity are recorded using surface electrodes, it does not necessarily mean asymmetrical changes in intracranial activity have taken place. However, the source analysis did support a left hemisphere origin for these changes and one possible interpretation for seeing enhanced activity over the left hemisphere for the VOT stimuli might relate to the acoustic/linguistic content of these trained sounds. For example, Grimm et al. 2006 have shown a left-hemispheric preponderance of temporal information processing; as have Zatorre and Belin (2001), who point to enhanced myelination in the primary auditory cortex of the left hemisphere, when compared to the right hemisphere, which may favor processing of temporal information. More specifically, the left hemisphere has been shown to

have enhanced sensitivity to encoding VOT (Trébuchon-Da Fonseca et al., 2005; Sandman et al. 2007). Despite this supportive evidence, using non-linguistic control stimuli in the future, as well as binaural stimulation, will enable us to better understand the observed laterality effects reported here.

We should also keep in mind that training exercises can involve stimulus exposure, attention, focused listening, memory, decision making, and task execution. In this respect, auditory learning can therefore result from enhanced top-down cognitive processing as well as bottom-up sensory processing (Moore and Amitay, 2007) and it is possible that neural mechanisms associated with one or more of these processes contribute to the post-training findings reported here. Sheehan and colleagues (2005), for example, suggest that increases in P2 amplitude result from repeated stimulus exposure, and are not necessarily related to training. While stimulus exposure is an essential component to training, there is evidence to suggest that exposure alone is insufficient to account for all of the training-related physiological changes reported in the literature. For example, numerous studies show good test-retest reliability for N1 and P2 responses suggesting that exposure to stimuli during one test session does not automatically alter the physiological representation of sound during a second test session (for a review see Tremblay, 2007). In addition, Alain and colleagues (2007) showed that repeated exposure to sound can *minimize*, rather than enhance, P2 responses.

Furthermore, despite being exposed to the same number of stimuli, not all individuals showed enhanced P2 amplitudes following training. This means that not all individuals are equally affected by whatever processes were activated during the test and training sessions used here. This finding is of interest because it suggests that the neural mechanisms underlying the observed changes are more labile in some people compared to others. As an example, an interesting and unexpected finding was that pre-training N1 amplitudes were smaller in people who showed the least amount of perceptual change. One could speculate that individuals with smaller N1 responses have auditory systems that are less synchronized to onset of sound, or less responsive to stimulus exposure. These explanations, however, still do not address why the training related changes were seen for P2 and not N1.

It is difficult to say exactly how N1 might contribute to the P2 findings presented in this study because little is known about the P2 response. Because P2 often co-varies with N1 along many stimulus dimensions, N1 and P2 are sometimes regarded to be subcomponents of a unitary response (e.g., N1-P2 peak-to-peak amplitude). However, there is also evidence to suggest there is some degree of independence between these two measures (Godey et al. 2001). For example we and others have reported abnormal P2 responses in the presence of normal N1 potentials in older adults with impaired speech understanding (Bertoli et al., 2002; Tremblay et al., 2003; Tremblay et al., 2004; Harkrider et al., 2005; Ross et al., 2007). And when combined with the effects of training on P2 reported here; especially in the absence of significant changes in N1, there is evidence to suggest that the functional significance of P2 might be underestimated.

Even though P2 amplitude increases were seen in some people who improved their ability to identify the stimuli used during training, and less evident in people who could be described as non-learners, there was no clear cut brain-behavior relationship involving P2, at least not when looking at time-locked evoked potentials recorded from midline, GFP, or source waveforms. But as previously suggested by us and others, because N1-P2 responses can be modulated by attention and other top-down processes (Hillyard et al., 1973; Woldorff and Hillyard, 1991; Alain and Woods, 1997), some of the physiological changes reported here may reflect top-down modulatory influences that take place while listening to the stimuli during evoked potential testing, or activated during the training task (Polley et al. 2006; Fritz et al. 2007). If enhanced P2 responses reflect heightened awareness or attention resulting from passive

listening during ERP recordings, we might expect to see similar stimulation patterns of brain activity for all of the stimuli being tested. Yet in our study, the distribution of P2 change was different for the control and experimental stimulus conditions. One interpretation is that enhanced P2 activity seen for all stimulus types and is therefore not stimulus specific, reflects general processes (such as arousal, awareness) that are activated during the experiment. In contrast, the lateralization effects seen only for the trained stimuli could have been shaped by task-dependent attention to the acoustic feature (VOT), activated during training. In this respect, focused attention to the voiced VOT cue could have contributed to enhanced temporal encoding in the left hemisphere, similar to the asymmetrical hemispheric specializations reported in humans, described earlier. What is more, this interpretation is also in line with some of the animal literature where neural correlates of task-dependent plasticity have been reported (Fritz, Elhilali, and Shamma, 2005; Polley, Steinberg, and Merzenich, 2006). For example, when rats were trained to attend to either frequency or intensity cues, topographical reorganizations corresponded only to the specific acoustic feature that was attended to (Polley et al. 2006).

Of course it is also possible that participating in this experiment, and learning to assign a linguistic category to ‘mba’ and ‘ba’, contributes to the lateralized representation of the experimental stimuli only. Training presumably makes the experimental stimuli more meaningful and more memorable than the control stimulus. The P2 has been described as reflecting multiple processes including analyzing acoustical features and the formation of auditory memory (Naatanen & Picton, 1987), and there is speculation that positive ERP components in the 200–250 ms latency range are related to stimulus identification (Crowley & Colrain, 2004).

In conclusion, it can be said that training tunes both bottom-up and top-down neural mechanisms. Some changes are likely specific to the trained stimulus and some reflect more generalized processing. Here we conclude by stating that there are stimulus-specific and non-specific effects of training that can be measured in humans, and that patterns of brain activity, as well as the presence or absence of brain-behavior relationships, can help define the underlying neural mechanisms affected by this type of training protocol. In our case, the effects of training differed between control and experimental stimuli. Whereas the effects of training on P2 amplitude were greatest in the left hemisphere for the trained stimuli, enhanced P2 activity was seen in both hemispheres for the control stimulus. We also demonstrate that some of the effects of training are lost when data analysis is limited to Cz or GFP. Stimulus specific effects were only seen when data from both hemispheres were compared. This point is important because investigators often limit their analyses to electrode site Cz or GFP, with the intention of translating these methods and findings to clinical situations when it is not feasible to use a large number of electrodes. Here we emphasize that the observed patterns of evoked activity seen across hemispheres, which differed for each stimulus type, would have been missed in a clinical situation had data from only one electrode been analyzed.

Acknowledgements

This work was supported by the National Institutes of Health [NIDCD R01 DC007705] awarded to KT; the Virginia Merrill Bloedel Hearing Research Traveling Scholar Program; and the National Institutes of Health [P30 DC04661] participant recruitment pool. Portions of this experiment were presented at the International Evoked Response Audiometry Study Group (IERASG) in Vancouver of 2001.

References

- Alain C, Woods DL. Attention modulates auditory pattern memory as indexed by event-related brain potentials. *Psychophysiology* 1997;34:534–546. [PubMed: 9299908]

- Alain C, Snyder JS, He Y, Reinke KS. Changes in Auditory Cortex Parallel Rapid Perceptual Learning. *Cereb Cortex* 2007;17:1074–1084. [PubMed: 16754653]
- Amitay S, Irwin A, Hawkey DJ, Cowan JA, Moore DR. A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners. *J Acoust Soc Am* 2006;119:1616–1625. [PubMed: 16583906]
- Atienza M, Cantero JL, Dominguez-Marin E. The time course of neural changes underlying auditory perceptual learning. *Learn Mem* 2002;9:138–150. [PubMed: 12075002]
- Barlow, HB.; Foldiak, P. Adaptation and decorrelation in the cortex. In: Miall, C.; Durbin, RM.; Mitchison, GJ., editors. *The Computing Neuron*. New York: Addison-Wesley; 1989. p. 54-72.
- Bertoli S, Smurzynski J, Probst R. Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task. *Clin Neurophysiol* 2002;113:396–406. [PubMed: 11897540]
- Bosnyak DJ, Eaton RA, Roberts LE. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cereb Cortex* 2004;14:1088–1099. [PubMed: 15115745]
- Crowley, Colrain Crowley KE, Colrain IM. A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol* 2004;115(4):732–744. [PubMed: 15003751]
- Dahmen JC, King AJ. Learning to hear: plasticity of auditory cortical processing. *Curr Opin Neurobiol* 2007;17:456–464. [PubMed: 17714932]
- Eggermont JJ. Representation of a voice onset time continuum in primary auditory cortex of the cat. *The Journal of the Acoustical Society of America* 1995;98:911–920. [PubMed: 7642830]
- Fritz JB, Elhilali M, David S, Shamma S. Auditory attention -- focusing on the searchlight in sound. *Curr Opin Neurobiol* 2007;17:437–455. [PubMed: 17714933]
- Gilbert CD, Sigman M, Crist RE. The neural basis of perceptual learning. *Neuron* 2001;31:681–697. [PubMed: 11567610]
- Godey B, Schwartz D, de Graaf JB, Chauvel P, Liégeois-Chauvel C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin Neurophysiol* 2001;112:1850–1859. [PubMed: 11595143]
- Grimm S, Roeber U, Trujillo-Barreto NJ, Schröger E. Mechanisms for detecting auditory temporal and spectral deviations operate over similar time windows but are divided differently between the two hemispheres. *Neuroimage* 2006;32:275–282. [PubMed: 16651011]
- Harkrider AW, Plyler PN, Hedrick MS. Effects of age and spectral shaping on perception and neural representation of stop consonant stimuli. *Clin Neurophysiol* 2005;116:2153–2164. [PubMed: 16043402]
- Harville, DA. *Matrix Algebra From a Statistician's Perspective*. Springer-Verlag; New York: 1997. p. 85-86.
- Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science* 1973;182:177–180. [PubMed: 4730062]
- Irvine DR, Wright BA. Plasticity of spectral processing. *Int Rev Neurobiol* 2005;70:435–472. [PubMed: 16472642]
- Klatt D. Software for cascade/parallel formant synthesizer. *J Acoust Soc Am* 1980;67:971–995.
- Lakshminarayanan K, Tallal P. Generalization of non-linguistic auditory perceptual training to syllable discrimination. *Restorative Neurology and Neuroscience* 2007;25:263–272. [PubMed: 17943004]
- MacMillan, NA.; Creelman, CD. *Detection theory: a user's guide*. Cambridge Univ Press; 1991.
- McClaskey CL, Pisoni DB, Carrell TD. Transfer of training of a new linguistic contrast in voicing. *Percept Psychophys* 1983;34:323–330. [PubMed: 6657433]
- Menning H, Imaizumi S, Zwislock P, Pantev C. Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrasts of the Japanese language. *Learn Mem* 2002;9:253–267. [PubMed: 12359835]
- Moore DR, Amitay S. Auditory Training: Rules and Applications. *Sem Hear* 2007;28:099–109.

- Naatanen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 1987;24(4):375–425. [PubMed: 3615753] 1987
- Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci* 2006;26:4970–4982. [PubMed: 16672673]
- Reinke KS, He Y, Wang C, Alain C. Perceptual learning modulates sensory evoked response during vowel segregation. *Brain research* 2003;17:781–791. [PubMed: 14561463]
- Ross B, Tremblay KL, Picton TW. Physiological detection of interaural phase differences. *J Acoust Soc Am* 2007;121:1017–1027. [PubMed: 17348524]
- Sandmann P, Eichele T, Specht K, Jancke L, Rimol LM, Nordby H, Hugdahl K. *Rest Neurol Neurosci* 2007;25:227–240.
- Scherg M. Fundamentals of dipole source potential analysis. *Adv Audiol* 1990;6:40–69.
- Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci* 2003;23:5545–5552. [PubMed: 12843255]
- Shahin A, Roberts LE, Pantev C, Trainor LJ, Ross B. Modulation of the P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport* 2005;16:1781–1785. [PubMed: 16237326]
- Sheehan KA, McArthur GM, Bishop DV. Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Brain Res Cogn Brain Res* 2005;25:547–553. [PubMed: 16198089]
- Skrandies, W. Topographical analysis of electrical brain activity: methodological aspects. In: Zani, A.; Mado Proverbio, A., editors. *The Cognitive Electrophysiology of Mind and Brain*. San Diego: 2003. p. 401-416.
- Steinschneider M, Schroeder CE, Arezzo JC, Vaughan HG Jr. Physiologic correlates of the voice onset time boundary in primary auditory cortex (A1) of the awake monkey: temporal response patterns. *Brain Lang* 1995;48:326–340. [PubMed: 7757450]
- Tallal P. Improving language and literacy is a matter of time. *Nat Rev Neurosci* 2004;5:721–728. [PubMed: 15322530]
- Trébouchon-Da Fonseca A, Giraud K, Baider JM, Chauvel P, Liégeois-Chauvel C. Hemispheric lateralization of voice onset time (VOT) comparison between depth and scalp EEG recording. *Neuroimage* 2005;27:1–14. [PubMed: 15896982]
- Tremblay KL. Training-related changes in the brain: Evidence from human auditory evoked potentials. *Sem Hear* 2007;28:120–132.
- Tremblay K, Ross B. Effects of age and age-related hearing loss on the brain. *J Commun Disord* 2007;40:305–312. [PubMed: 17433356]
- Tremblay K, Kraus N, McGee T. The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport* 1998;9:3557–3560. [PubMed: 9858359]
- Tremblay K, Kraus N, Carrell TD, McGee T. Central auditory system plasticity: generalization to novel stimuli following listening training. *J Acoust Soc Am* 1997;102:3762–3773. [PubMed: 9407668]
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear* 2001;22:79–90. [PubMed: 11324846]
- Tremblay KL, Kraus N. Auditory training induces asymmetrical changes in cortical neural activity. *J Speech Lang Hear Res* 2002;45:564–572. [PubMed: 12069008]
- Tremblay KL, Piskosz M, Souza P. Effects of age and age-related hearing loss on the neural representation of speech cues. *Clin Neurophysiol* 2003;114:1332–1343. [PubMed: 12842732]
- Tremblay KL, Billings C, Rohila N. Speech evoked cortical potentials: effects of age and stimulus presentation rate. *J Am Acad Audiol* 2004;15:226–237. [PubMed: 15119463]
- Woldorff MG, Hillyard SA. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 1991;79:170–191. [PubMed: 1714809]
- Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 2001;11:946–953. [PubMed: 11549617]

Figure 1. Two pre-voiced stimuli illustrated as time waveforms and spectrographs. The 'mba' stimulus has 20 ms of pre-voicing, and the 'ba' stimulus has 10 ms of pre-voicing (shaded areas).

Figure 2.
Flowchart describing procedure and stimulus conditions.

Figure 3. Performance significantly improved following training. Group and individual d-prime values (+/- std error bars) are shown. 'Non-learners' are indicated as 'n'.

Figure 4. Group averaged pre- (thin line) and post-training (thick line) P1-N1-P2 waveforms (n=13). Significant increases in P2 amplitude are evident in each stimulus condition.

Figure 5. Pre- (thin line) and post-training GFP measures for each stimulus condition 'mba', 'ba' and 'a'. Considerable increases in P2 amplitude are seen for all stimulus conditions following training.

Figure 6.

A) Regional source consisting of three orthogonal vectors: R=radial (medial-lateral), T1=first tangential (inferior-superior), and T2= second tangential (anterior-posterior) components. B) Pre- and post-training P1-N1-P2 source waveforms for each stimulus, hemisphere and source [tangential1 (T1), radial (R), and tangential2 (T2)]. For the control stimulus, increases in P2 are apparent over both hemispheres in the T1 condition. For the experimental stimuli, P2 changes are greater over the left hemisphere.

Figure 7.

A Individuals with smaller pre-training N1 amplitudes showed the least improvement in perception (measured in d-prime) with training. N1 amplitude values for each stimulus ($\circ$ = 'mba' and $\bullet$ = 'ba'), for each individual, are joined with a line (----). B) P1-N1-P2 recordings from electrode Cz for learners (n=10) and non-learners (n=3) in response to the stimulus 'ba'. N1 amplitude peaks are smaller for the non-learners, compared to the learners, pre- (black) and post- (red) training. The strength of N1 is also shown in voltage maps where the strength of N1 can be seen in blue on the top of the head.