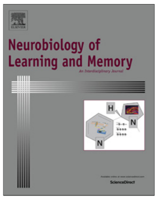
Neurobiology of Learning and Memory 109 (2014) 82–93

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme



Human brainstem plasticity: The interaction of stimulus probability and auditory learning

Erika Skoe [a,d,1], Bharath Chandrasekaran [b], Emily R. Spitzer [a,d], Patrick C.M. Wong [c,d,f], Nina Kraus [a,d,e,g,*]

[a] Auditory Neuroscience Laboratory, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA
[b] Department of Communication Sciences and Disorders, Center for Perceptual Systems, Institute for Neuroscience, The University of Texas at Austin, Austin, TX, USA
[c] Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong Special Administrative Region
[d] Department of Communication Sciences & Disorders, School of Communication, Northwestern University, Evanston, IL 60208, USA
[e] Department of Otolaryngology – Head and Neck Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA
[f] Department of Neurobiology & Physiology, Northwestern University
[g] Department of Otolaryngology, Northwestern University

ARTICLE INFO

Article history:
Received 1 June 2013
Revised 14 October 2013
Accepted 18 November 2013
Available online 26 November 2013

Keywords:
Neuroplasticity
Brainstem
Auditory learning

ABSTRACT

Two forms of brainstem plasticity are known to occur: an immediate stimulus probability-based and learning-dependent plasticity. Whether these kinds of plasticity interact is unknown. We examined this question in a training experiment involving three phases: (1) an initial baseline measurement, (2) a 9-session training paradigm, and (3) a retest measurement. At the outset of the experiment, auditory brainstem responses (ABR) were recorded to two unfamiliar pitch patterns presented in an oddball paradigm. Then half the participants underwent sound-to-meaning training where they learned to match these pitch patterns to novel words, with the remaining participants serving as controls who received no auditory training. Nine days after the baseline measurement, the pitch patterns were re-presented to all participants using the same oddball paradigm. Analysis of the baseline recordings revealed an effect of probability: when a sound was presented infrequently, the pitch contour was represented less accurately in the ABR than when it was presented frequently. After training, pitch tracking was more accurate for infrequent sounds, particularly for the pitch pattern that was encoded more poorly pre-training. However, the control group was stable over the same interval. Our results provide evidence that probability-based and learning-dependent plasticity interact in the brainstem.

1. Introduction

Natural auditory environments are complex and nonstationary. This poses a challenge for the nervous system, which must find structure and meaning within the ever-changing soundscape. The computational complexity of this task is masked by the near seamless way in which most humans are able to adapt "on the fly" to a great variety of natural environments. To keep pace with the auditory landscape, the brain must accommodate minute changes in the environment while at the same time tracking, cataloging, and grouping repetitive sounds based on their behavioral relevance. This computational dexterity, which can be observed in all sensory systems, allows the brain to learn a great deal of information about the environment, both in the immediate and long run. Through continued experience with the world, this process is refined and altered, leading to a complex interaction between on-line and past learning, with the ability to learn in the moment reflecting the collective experiences of the individual (Herholz, Boh, & Pantev, 2011).

This study examines the interaction between on-line probability detection and auditory learning within the human auditory brainstem. Until now, these forms of brainstem plasticity have been explored independently in animal models (Gao & Suga, 1998, 2000; Malmierca, Cristaudo, Perez-Gonzalez, & Covey, 2009; Malone & Semple, 2001; Perez-Gonzalez, Malmierca, & Covey, 2005; Suga, 2008; Suga, Gao, Zhang, Ma, & Olsen, 2000; Suga, Xiao, Ma, & Ji, 2002; Szymanski, Garcia-Lazaro, & Schnupp, 2009) and humans. In humans, short-term auditory training paradigms—in which

---

Abbreviations: ABR, auditory brainstem response; FFR, frequency-following response; IC, inferior colliculus; SSA, stimulus-specific adaptation.

* Corresponding author.

E-mail addresses: erika.skoe@uconn.edu (E. Skoe), bchandra@austin.utexas.edu (B. Chandrasekaran), emilyspitzer2013@u.northwestern.edu (E.R. Spitzer), p.wong@cuhk.edu.hk (P.C.M. Wong), nkraus@northwestern.edu (N. Kraus).

URLs: http://skoe.slhs.uconn.edu (E. Skoe), http://soundbrainlab.wordpress.com/ (B. Chandrasekaran), http://www.brainvolts.northwestern.edu (E.R. Spitzer), http://brain.cuhk.edu.hk/ (P.C.M. Wong), http://www.brainvolts.northwestern.edu (N. Kraus).

[1] Present address: Department of Speech, Language, and Hearing Sciences, Department of Psychology Affiliate, Cognitive Science Program Affiliate, University of Connecticut, 850 Bolton Rd., Storrs, CT 06269, USA.

1074-7427/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.nlm.2013.11.011

participants learn new sound contrasts or are trained to hone their communication skills—have revealed that the auditory brainstem undergoes learning-dependent changes over the course of several weeks (Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Carcagno & Plack, 2011; Chandrasekaran, Kraus, & Wong, 2012; Song, Skoe, Wong, & Kraus, 2008). A separate line of research has revealed that the human auditory brainstem is sensitive to stimulus statistics, including the probability with which a sound is presented during an experiment (Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Parbery-Clark, Strait, & Kraus, 2011; Skoe & Kraus, 2010b; Skoe, Krizman, Spitzer, & Kraus, 2013; Slabu, Grimm, & Escera, 2012). As an example, Chandrasekaran and colleagues found that brainstem activity is enhanced when the stimulus train contains a single (predictable) sound compared to a pseudo-randomized mix of sounds in which the target sound is presented both infrequently and unpredictably (Chandrasekaran et al., 2009). However, whether probability-dependent processes in the brainstem can be changed through short-term training remains an unanswered question.

To examine these forms of plasticity in the human auditory brainstem and how they interact, we combined an oddball paradigm with a short-term auditory training paradigm. Oddball paradigms are routinely used to assess human cortical change detection (e.g., Kirmse et al., 2008; Kujala & Naatanen, 2010; Naatanen, 1995, 2008; Naatanen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001) and more recently they have been adopted for studying human subcortical change detection (Slabu, Escera, Grimm, & Costa-Faidella, 2010; Slabu et al., 2012). We noninvasively measured far-field auditory brainstem responses (ABRs) to time-varying pitch patterns using electrodes placed on the scalp (Skoe & Kraus, 2010a). Analyses focused on the tonic component of the ABR known as the frequency-following response (FFR). The FFR, which reflects phase-locking from neuronal ensembles within the auditory brainstem and midbrain (Akhoun, Gallego, Moulin, Menard, Veuillet, Berger-Vachon, Collet, and Thai-Van, 2008; Akhoun, Moulin, Jeanvoine, Menard, Buret, Vollaire, Scorretti, Veuillet, Berger-Vachon, Collet, and Thai-Van, 2008; Chandrasekaran & Kraus, 2010; Chandrasekaran et al., 2009), captures complex spectrotemporal information in speech, such as changes in fundamental frequency (the physical correlate of pitch), with a high degree of precision (Krishnan, Xu, Gandour, & Cariani, 2004, 2005).

This study involved three phases: (1) an initial baseline measurement of auditory brainstem activity, (2) a 9-session auditory training paradigm, and (3) retest measurement of auditory brainstem activity (Fig. 1). In phases 1 and 3, complex pitch patterns (Mandarin tones 1 and 4, "T1" and "T4") were each presented in two different probability conditions. In one condition, the pitch pattern was presented with a high probability of occurrence (i.e., "standard") and in the other (reversed) condition it was presented with a low probability of occurrence (i.e., "deviant"). In tonal languages such as Mandarin, these pitch patterns are suprasegmental cues that carry lexical information within a syllable. These pitch patterns, however, are not linguistically meaningful to (untrained) monolingual English-speaking participants. The goal of the training paradigm was to make these pitch patterns linguistically meaningful. Half of the participants were trained to use these unfamiliar pitch patterns to distinguish words, the other half served as the control group. The control group was retested at an interval equivalent to that of the trained participant but received no additional experimental stimulation between the baseline and retest sessions.

By studying the effects of auditory training on the FFR to the same acoustic token, we are able to examine the extent to which the auditory brainstem is modulated by short-term auditory learning. Consistent with recent reports using similar auditory training paradigms, we predicted that pitch patterns would be tracked more precisely after training (Chandrasekaran et al., 2012; Song et al., 2008). In addition, by presenting the same acoustic token under standard and deviant conditions, we are able to examine the effects of stimulus probability on brainstem pitch tracking. Previous investigations have largely focused on how the magnitude of the FFR is changed when the stimulus probability changes (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Skoe & Kraus, 2010b; Skoe et al., 2013; Slabu et al., 2012). Here we focus on a different, more fine-grained aspect of the FFR, namely the precision with which the stimulus pitch contour is tracked when the stimulus is presented frequently or infrequently within a stimulus train. Given that stimulus predictability is associated with enhanced FFR magnitudes for sounds with simple (flat) pitch contours (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Slabu et al., 2012), we likewise predicted that pitch tracking would be more precise for the standard relative to the deviant condition. By combining these two types of experimental paradigms we were able to examine whether probability-detection mechanisms in the brainstem are changed when the stimulus acquires new meaning to the listener. Drawing on work from Parbery-Clark et al. (2011), we hypothesized that the neural mechanisms that underlie on-line probability detection can be changed through experience-dependent processes such as auditory learning. Parbery-Clark et al. demonstrated that highly trained musicians (with an average of 16 years of training) have FFR enhancements (i.e., larger response to the fundamental frequency of the stimulus) to statistically-probable sounds versus rare sounds (Parbery-Clark et al., 2011). This is in contrast to age-matched non-musicians in whom frequent and rare sounds did not elicit different FFR outcomes. We predicted that training would change probability-dependent activity, resulting in one of two distinct outcomes. One predicted outcome was that probability-dependent processes would become heightened such that responses to the standard and deviant

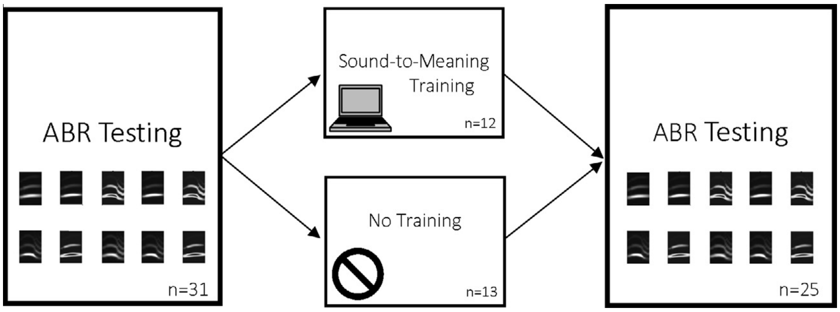


Fig. 1. Experimental design. Auditory brainstem responses (ABRs) were recorded on 31 participants using an oddball design (Fig. 2) as part of an initial session of baseline measurements. Participants were then split into two groups: 12 underwent 9 sessions of sound-to-meaning training and 13 received no training (6 participants did not continue with the experiment). After 9 days, the trained and control groups were retested using the same oddball paradigm.

conditions would become more different after the listener had gained practice utilizing these pitch contours within a linguistic context to distinguish words. That is, by becoming more sensitive to these contrasts, the listener would be more sensitive to changes in stimulus probability and therefore have larger differences in pitch tracking between standard and deviant conditions after training. The second predicted outcome was that increased familiarity to pitch contours, coupled with their changed linguistic significance, would minimize differences between the standard and deviant conditions, such that after training deviant sounds would sound less deviant and give rise to more accurate pitch tracking.

2. Materials and methods

2.1. Participants

All study participants were native speakers of American English with no prior exposure to tone languages. Participants had a range of past musical training (<6 years of continuous musical training, not starting before the age of 7 years) but were not currently practicing. Previous studies have shown that on-going musical training modulates the FFR to non-native pitch patterns (Bidelman, Gandour, & Krishnan, 2011; Wong, Skoe, Russo, Dees, & Kraus, 2007), hence participants with significant musical expertise were excluded. Participants had normal hearing defined as air conduction thresholds <20 dB nHL for octaves from 250 to 8000 Hz. Consistent with Northwestern University Institution Review Board procedures, informed consent was obtained from all participants. All electrophysiological testing took place in the Auditory Neuroscience Laboratory at Northwestern University.

Thirty-one young adults participated in phase 1 of the experiment, during which we obtained baseline ABR measurements (phase 1) (mean age 26 ± 4.92 years, 10 females). In phase 2, participants were pseudo-randomly assigned to one of two groups: auditory training ($n = 18$) or control ($n = 13$). As we neared the end of study enrollment, recruitment was targeted to ensure that the two groups were matched on sex and age. Twelve of the trained participants and 13 of the control participants returned to the laboratory to undergo electrophysiological retesting (phase 3). Six of the trained participants did not complete phase 3, which occurred 9 days (on average) after the initial baseline measurements. For phases 2 and 3, data are reported from 12 trained and 13 control, who were matched on the group level in terms of age ($t(23) = 1.39$, $p = 0.18$) and sex (Kolmogorov-Smirnov $Z = 0.201$, $p = 1.0$).

2.2. Electrophysiological testing

2.2.1. Stimuli

In phase 1 and phase 3 of the experiment, ABRs were recorded to two 278.5 ms syllables (/T1/ and /T4/). Both stimuli are composed of the same syllable, /mi/, produced by a native speaker of Mandarin. The stimuli differ on their pitch trajectory: /T1/ has a relatively level pitch trajectory that rises (non-linearly) from 143 to 171 Hz and /T4/, in comparison, has a falling pitch trajectory that starts at 148 Hz and midway through the syllable drops non-linearly to 94 Hz (Fig. 2). These stimuli were identical to ones used in previous experiments (Song et al., 2008; Wong et al., 2007). In comparison to the other two Mandarin contours (/T2/ (rising pitch) and /T3/ (dipping pitch), FFR pitch tracking is weaker for T1 and T4 (Krishnan et al., 2004).

2.2.2. Oddball paradigm

To examine the effects of stimulus probability on auditory brainstem responses, T1 and T4 were presented in an oddball paradigm Fig. 2. In one condition, T1 served as the standard (presented 80% of the time) and T4 as the deviant (presented 20% of the time) (T1s/T4d). In the other condition, they were reversed so that T4 was the standard (80%) and T1 was the deviant (20%) (T4s/T1d). The order of the conditions was randomized across participants. Comparing standard vs. deviant responses to the same acoustic token, a method used in examining cortical change detection (Jacobsen & Schroger, 2003; Jacobsen, Schroger, Horenkamp, & Winkler, 2003), controls for acoustical differences between stimuli while changing only the probability of occurrence.

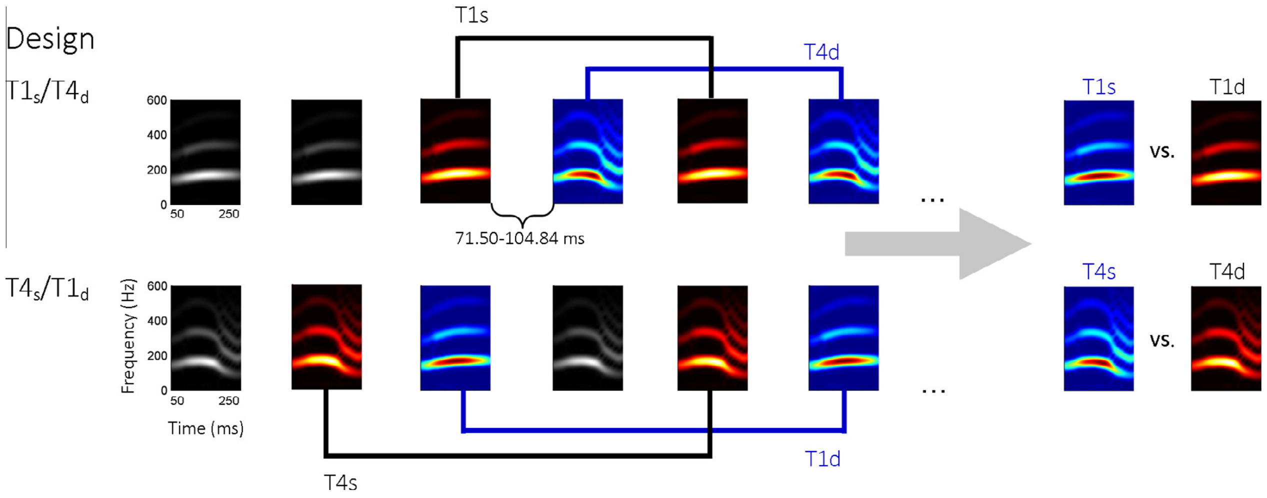


Fig. 2. An oddball design used to elicit auditory brainstem responses (ABR) to pitch patterns. The stimulus set included a rising (T1) and a falling (T4) pitch pattern. T1 and T4 were presented together under two different conditions. In one condition (schematized in the top row), T1 was presented frequently ("standard", 80%), and T4 was presented rarely ("deviant", 20%). In the other condition (bottom), the pattern was reversed; T4 was presented as the standard and T1 as the deviant. In this schematic, the standard tones are plotted using either grayscale or red-colorized spectrograms (time–frequency plots) and the deviant tones are plotted in a blue spectrograms. Stimuli were presented at a randomized inter-stimulus interval that ranged from 71.50 to 104.84 ms. To examine the effect of stimulus probability on the ABR, responses to T1 presented as the standard stimulus (T1s) were compared with responses to T1 presented as deviant (T1d), and similarly responses to T4s were compared with T4d. To ensure that an equivalent number of standard and deviant trials were compared, the trials immediately preceding the deviants (red spectrogram) were selected to represent the response to standard condition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2.3. Passive ABR recording

During electrophysiological testing, participants sat in an acoustically and electrically shielded chamber and were instructed to ignore the sounds presented binaurally via insert earphones (ER-3; Etymotic Research, Elk Grove Village, IL) at 75 dB SPL. Participants watched a movie of their choice (volume was muted and English subtitles were provided), a common practice for minimizing myogenic artifacts during electrophysiological recordings (Skoe & Kraus, 2010a). Stimuli were presented with a variable inter-stimulus interval between 71.50 and 104.84 ms. Responses were recorded with a sampling rate of 20 kHz using Scan 4.3 (Neuroscan; Compumedics Inc.) with Ag–AgCl scalp electrodes, recorded from Cz (active) to earlobes (reference), with the forehead as ground. Contact impedance was less than 5 kOhms for all electrodes. 3500 sweeps were collected per sequence (half of each polarity), resulting in 2800 sweeps for the standard tone and 700 sweeps for the deviant tone. The T4s/T1d and T1s/T4d conditions each lasted roughly 20 min.

2.2.4. Deriving responses to standard and deviant conditions

After band-pass filtering the continuous recording from 80 to 1000 Hz (12 dB/octave) in Edit 4.3 (Neuroscan; Compumedics Inc), ABRs to the deviant stimulus were averaged. To derive the average response to the standard condition, responses to the standard stimulus immediately before each deviant were averaged. Averaging occurred over a time window encompassing −45 ms prior to stimulus onset and 16.5 ms after stimulus offset. After rejecting trials with activity greater than ±35 μV, the average number of trials per condition was approximately 650.

2.2.5. Pitch tracking procedure

Analysis focused on how accurately the pitch contour of the stimulus was tracked in the ABR, and how this tracking changed as a function of stimulus probability and training. The pitch contours of the stimulus and response were derived using a sliding window (60-ms) autocorrelation-based procedure (Krishnan et al., 2004, 2005; Wong et al., 2007). This procedure was applied separately to the stimulus and response waveforms. To estimate how the pitch period changed throughout the waveform, the waveforms were divided into 219 bins, each 60 ms (59 ms overlap between adjacent time bins). Each 60-ms bin contained ∼9 cycles of the pitch periods of T1 and ∼7.5 cycles of T4, from which the period (i.e., the period over which the waveform repeats) was objectively estimated using an autocorrelation-based procedure. This procedure involved cross-correlating each bin with itself (autocorrelation) and finding the lag corresponding to the maximum autocorrelation. This point of maximum correlation was taken to be the pitch period of each bin. The maximum autocorrelation was found over a lag range of 3.33–14.3 ms, a range that encompasses the time-variant (respective) periods of the T1 and T4 pitch contours. The lag value associated with the maximum correlation was converted to frequency by taking its reciprocal (frequency = 1/period (s)). The resulting frequency values were strung together to form a 219-point pitch contour. This procedure is illustrated in Fig. 3.

The first bin began at time 0 ms for the stimuli (bin_{S1}: 0–60 ms, bin_{S2} = 1–61 ms, bin_{S219} = 218–278 ms). To account for the transmission delay of the ER-3 ear inserts as well as the neural delay, the analysis bin began at 10 ms for the responses (bin_{R1}: 10–70 ms, bin_{R2} = 11–71 ms, bin_{R219} = 228–288 ms). By starting the analysis at 10 ms, this also effectively removed the phasic component of the ABR, leaving the tonic FFR.

The primary dependent variable in this study is *Frequency Error*. *Frequency Error* of the response was derived by first doing a point-by-point subtraction between the stimulus and the response pitch contours, then taking the absolute value of the difference, and finally averaging across all 219 points. This resulted in a single num-

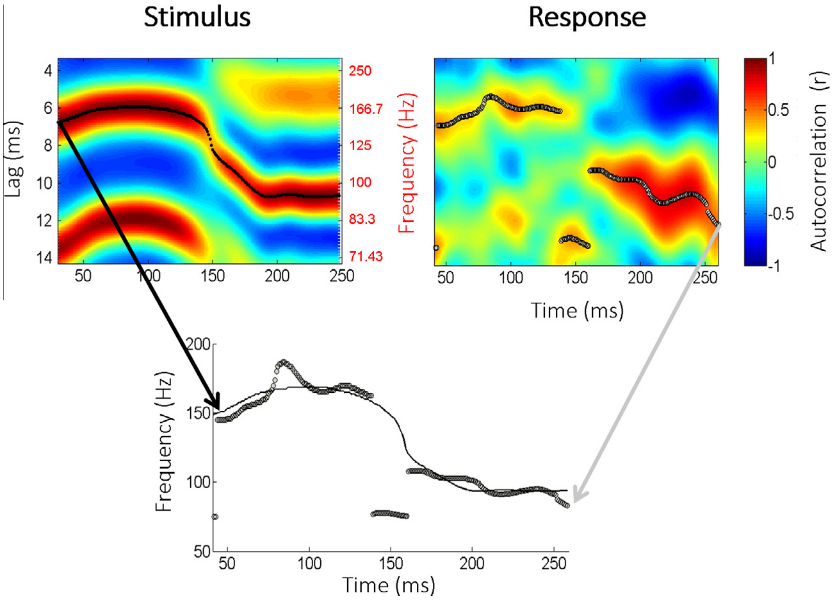


Fig. 3. Illustration of the *Frequency Error* measure. The *Frequency Error* measure, which serves as the primary dependent variable in this study, was derived by comparing the pitch contours of the stimulus and response. In this example, the T4 stimulus is plotted along with the response from a single participant. In the pitch contour plot (bottom), the stimulus contour is plotted in black and the response contour is plotted as gray circles. Each circle (219 in total) represents a separate analysis bin. In this Figure, as well as Figs. 5 and 6, each bin is plotted relative to its midpoint. Thus, the first bin of the response, which runs from 10 to 70 ms, is plotted at a time point of 40 ms. The first bin of the stimulus extends from 0 to 60 ms; however, here it has been plotted to align with the first bin of the response. The pitch contours of the stimulus and response were derived from their respective autocorrelogram (top). For each bin, the maximum autocorrelation was found. In this figure, it presents the point with the reddest color for each vertical slice of the autocorrelogram. For the purposes of demonstration, black dots also appear over the autocorrelation maxima for the stimulus, and gray circles appear over the autocorrelation maxima for the response. The lag corresponding to each maxima was converted to frequency (frequency = 1/lag (s)) and then plotted as a function of time to derive the pitch contour at the bottom. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ber reflecting how many Hz the response contour deviated, on average, from the stimulus contour.

2.3. Sound-to-meaning auditory training

The participant's task was to learn 24 words (6 syllables × 4 lexical tones) produced by four different talkers over nine training sessions. Procedures were identical to those from a previous study (Chandrasekaran et al., 2012). In each training session, the participant learned to pair sounds with objects represented by line drawings (e.g., /peshT1/ = "COW"; /peshT4/ = "SHOE") (Fig. 4). Each training session contained the following sequence of events: first, participants listened to each sound paired with the visual presentation of the object. Following the auditory-visual pairing, a short quiz was administered to test sound-to-object association skills. Feedback (correct/incorrect) was provided to facilitate correct sound-object pairings. At the end of each training session, a final test was conducted during which no feedback was provided. In this test, participants heard a word (e.g., /peshT4/) and had to pick one of 24 objects that referred to this word. In total, 96 sounds (24 × 4 talkers) were heard. A final score (proportion word identification) was calculated from each of the nine sessions (Fig. 4). To make the learning task more challenging than previous studies, this study used four pitch contours, instead of three, in addition to a multi-talker, instead of single talker paradigm (Song et al., 2008) (Fig. 4A).

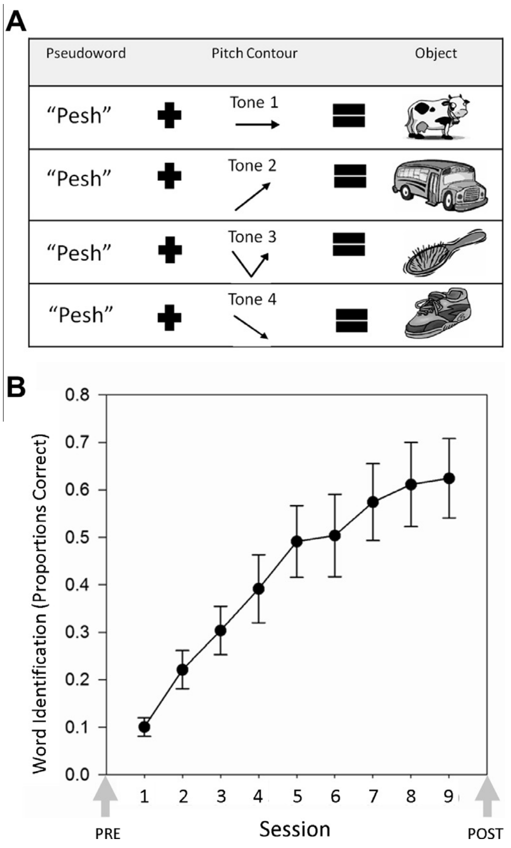


Fig. 4. Sound-to-meaning linguistic training. Participants ($n = 12$) underwent nine sessions of a sound-to-meaning training program during which they learned to use pitch patterns to distinguish 24 words. (A) Participants learned to associate six different pseudo words (pesh, dri, ner, ves, nuk, fjut), superimposed with one of four different pitch patterns (T1 = level tone, T2 = rising tone, T3 = dipping tone, T4 = falling tone) with 24 different objects. (B) The proportion of words correctly identified in each session is plotted as a function of training session. Participants showed significant learning across training sessions 1–5, after which they reached a plateau in word identification performance. Auditory brainstem responses to pitch patterns T1 and T4 were collected before (pre-training) and after training (post-training) (indicated by gray arrows) using procedures described in Figs. 2 and 3.

To construct the stimuli used for training, Mandarin pitch patterns produced by native Mandarin speakers (four male, four female) were superimposed onto six syllables (pesh, dri, ner, ves, nuk, fjut) produced by native American English speakers (four male, four female), resulting in 24 words minimally contrasted by pitch. The four Mandarin pitch patterns were: T1, T2, T3, and T4. The pitch patterns T1 and T4 were identical between the electrophysiological and training components of this study; however, the syllable used for evoking ABRs (/mi/) was not included among the set of trained syllables and it was produced by a different male talker than the set of trained syllables.

2.4. Statistics

For the baseline ABR measurements, *Frequency Error* was analyzed using a repeated measures ANOVA (RMANOVA), using Tone (T1 vs. T4) and Probability (standard vs. deviant) as independent variables. For the retest ABR measurement, a mixed-model RMANOVA was used to analyze *Frequency Error*, with 2 between-participant factors ((1) Group: trained vs. control and (2) test Session: baseline vs. retest) and 2 within participant factors ((1) Tone: T1 vs. T4 and (2) Probability: standard vs. deviant). Post-hoc pairwise comparisons were performed on significant effects.

3. Results

3.1. Phase 1: Baseline ABR measurements

Pitch tracking was more accurate (i.e., lower *Frequency Error*) for the standard condition vs. deviant condition (main effect of Probability: $F(1, 30) = 8.297$, $p = 0.007$; mean ± 1 standard deviation; standard: 34.17 ± 8.51 Hz vs. deviant: 37.09 ± 7.67 Hz). Moreover, T1 was tracked less precisely than T4 regardless of Probability, as evidenced by a main effect of Tone ($F(1, 30) = 127.706$, $p < 0.005$) but no Tone by Probability interaction ($F(1, 30) = 2.168$, $p = 0.151$) (Tone 1: 46.95 ± 11.00 Hz; Tone 4: 24.30 ± 7.52 Hz) (Fig. 5).

3.2. Phase 2: Sound-to-meaning training

After the initial electrophysiological testing, half of the participants underwent a nine session training program during which they learned to associate pitch patterns with words (Fig. 4). At the end of each training session, participants completed a word identification task. Scores from the word identification task obtained from each of the nine sessions (S1–S9) were entered into a one-way RMANOVA. The RMANOVA revealed a main effect of training ($F(1, 12) = 24.92$, $p < 0.001$). On average, participants showed significant learning from S1 to S5, and then a plateau in word-learning performance (Fig. 4). Statistically, word identification scores were significantly different for S5 relative to S1 ($t(12) = 7.38$, $p < 0.001$), but not S9 ($t(12) = 2.60$, $p = 0.15$), suggesting that learning stabilized after the fifth session (Fig. 4B).

3.3. Phase 3: Retest ABR measurements

We found a significant four-way interaction between Group, Test, Tone, and Probability ($F(1, 1, 1, 1) = 12.88$, $p = 0.002$) that was examined further using post-hoc comparisons exploring each component of the interaction.

Focusing first on the effects of training: At the initial baseline measurement, the groups were matched (no main effect of Group:

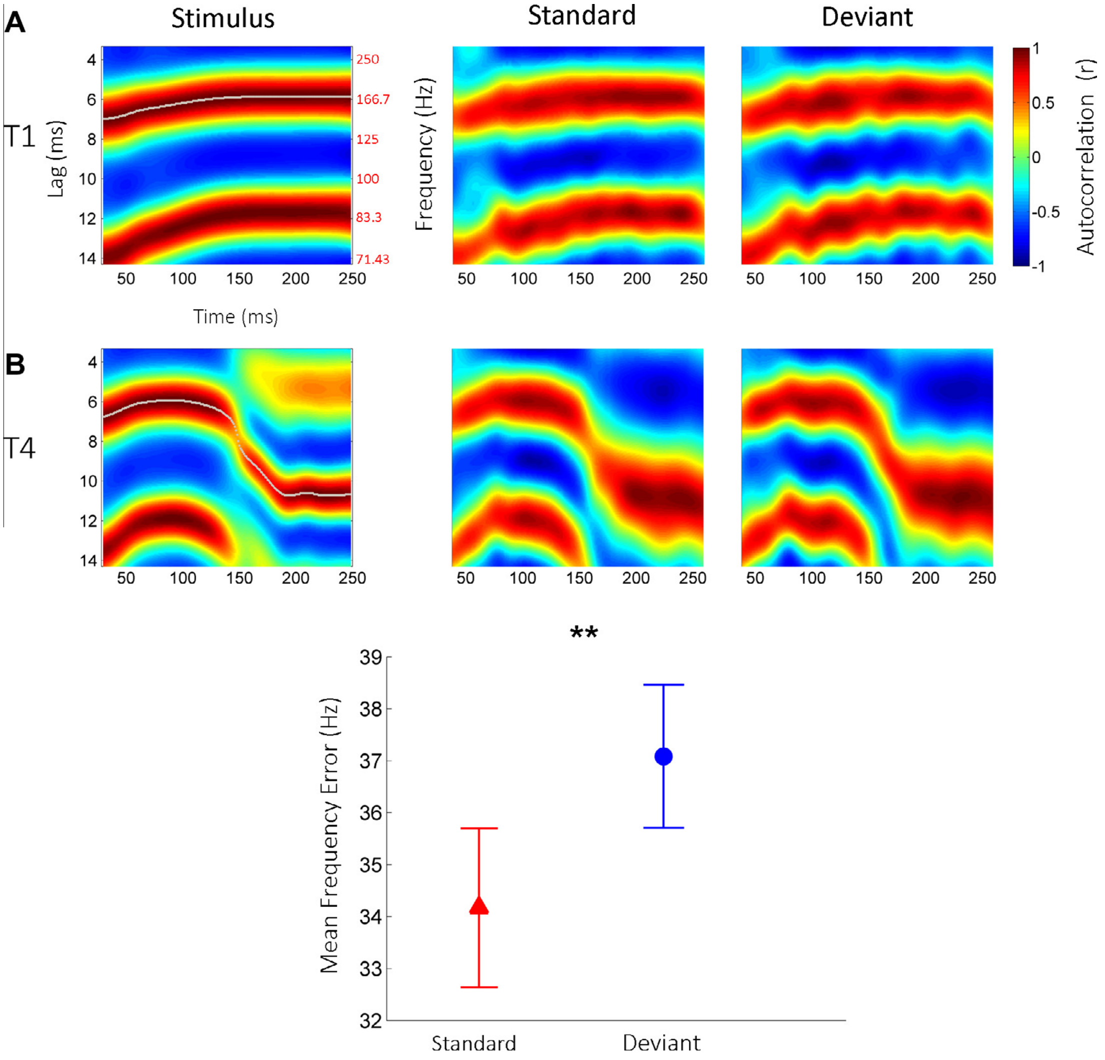


Fig. 5. Tracking of the stimulus pitch contour was more precise for the standard vs. deviant condition in the initial baseline measurement. (A) Autocorrelograms are plotted for T1 (top) and T4 (bottom). Color indicates the magnitude of the autocorrelation, with warmer colors signifying higher magnitudes than cooler colors. The standard condition appears in the middle, flanked by the deviant condition on the right and the stimulus on the left. In the stimulus panel, the thin white line represents the pitch contour of the stimulus. Pitching tracking to the stimulus was more accurate for the standard condition (middle) relative to the deviant condition (right). Compared to the deviant condition (right), the standard condition has a smoother-looking autocorrelogram that reflects more precise phase-locking to the stimulus contour and translates into lower *Frequency Errors*. The response autocorrelograms plotted here were derived from the time-domain group average waveforms. (B) Pitch tracking was more accurate (i.e., lower *Frequency Error*) for the standard condition vs. deviant condition (main effect of Probability: F(1, 30) = 8.297, $p = 0.007$). *Frequency Error* is plotted for the standard (black) and deviant (gray) conditions. Error bars signify 1 standard error from the mean. Because there was no interaction between Tone and Probability (F(1, 30) = 2.168, $p = 0.151$), the average *Frequency Error* across T1 and T4 is plotted. Asterisks indicate the level of significance ($^{**}p < 0.01$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(F(2, 23) = 0.35, $p = 0.60$; no Group × Probability interaction (F(2, 23) = 0.24, $p = 0.62$). However, over the course of the 9 day test-retest interval, pitch tracking improved for the trained but not for the control group, such that there was a main effect of Session in the trained group (F(1, 11) = 10.05, $p = 0.009$) that was not found in the control group (F(1, 12) = 0.54, $p = 0.48$). Collapsing across tones and probability conditions, the average *Frequency Error* dropped from 36.546 ± 24.42 to 32.293 ± 19.56 Hz for the trained group, but was more stable in the control group, starting initially at 34.849 ± 26.065 Hz and being 33.154 ± 25.220 Hz at retest.

Now examining the interaction between training and probability: In the initial test session, both groups individually showed a main effect of Probability (Fig. 6) (Trained: F(1, 11) = 4.766, $p = 0.040$; trending towards significance in the Control: F(1, 12) = 3.969, $p = 0.07$) and when combining the groups, the main effect of Probability was strong (F(1, 23) = 9.51, $p = 0.005$). However, at retest, in the trained group there was no overall effect of Probability (F(1, 12) = 0.656, $p = 0.435$) but instead a significant Tone × Probability interaction emerged (F(1, 111) = 10.542, $p = 0.008$). Interestingly, after training, *Frequency Error* for T1 was now larger for the standard condition relative to the deviant condition (standard *Frequency Error*: 47.99 ± 3.275 Hz; deviant Frequency Error = 38.208 ± 3.765 Hz, t(11) = 2.314, $p = 0.041$). For T4, pitch tracking was more equated between the standard and deviant conditions after training (standard Frequency Error: 18.911 ± 6.56 Hz; deviant Frequency Error = 24.064 ± 7.80 Hz, t(11) = −1.716, p = 0.114). Thus, the effects of training were most pronounced for T1 when it was a deviant sound, suggesting a complex interaction between training, tone, and probability (3-way interaction between Session, Tone, Probability: F(1, 1, 1) = 25.018, $p = 0.0004$). At the initial test session, the average Frequency Error for T1-deviants was 52.88 ± 8.66 Hz for the trained group and at retest it had fallen to 38.21 ± 13.04 Hz (t(11) = 4.29, $p = 0.001$). In contrast, for T1-standards, the trend was for tracking to get slightly

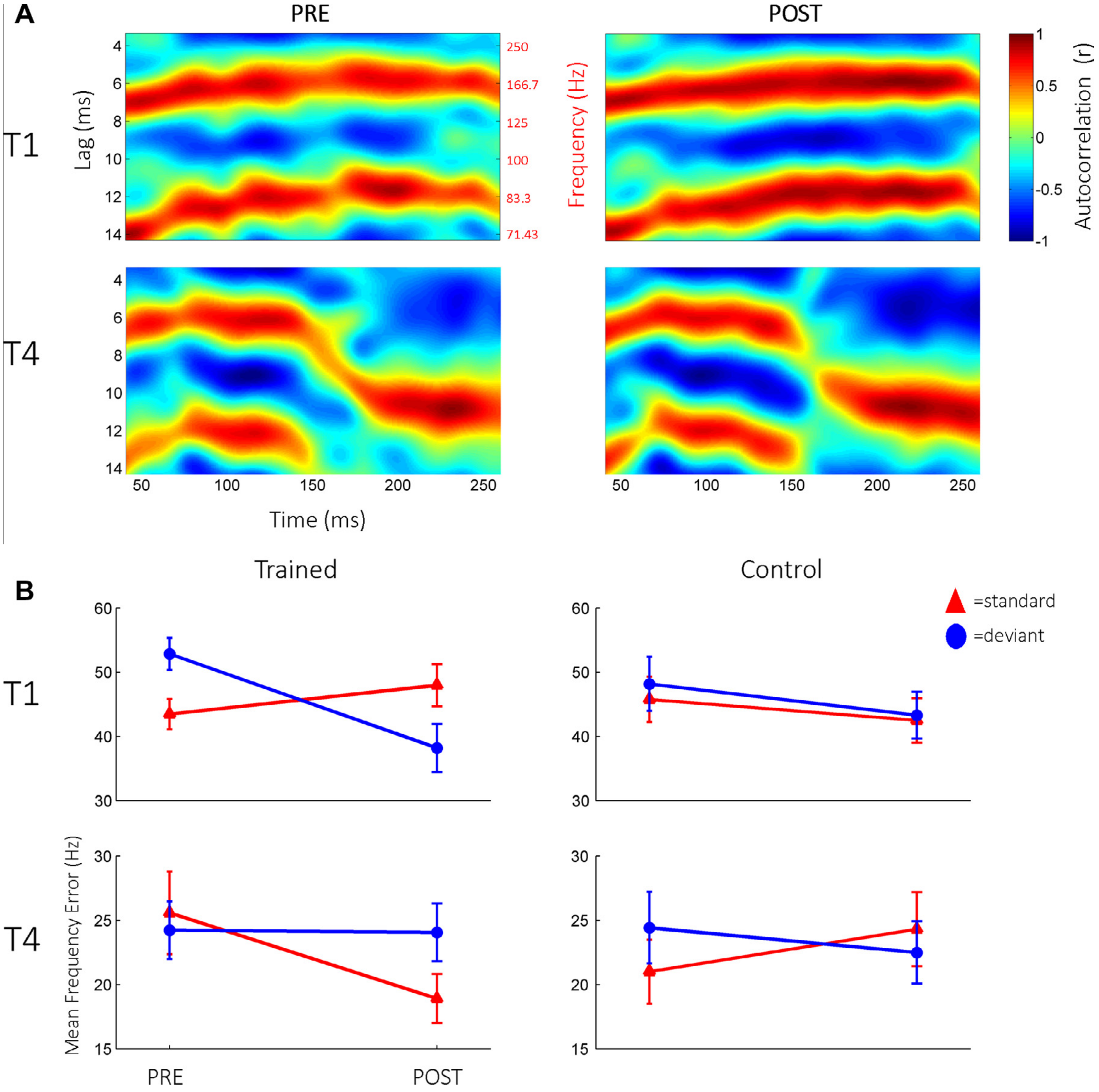


Fig. 6. Training is associated with better pitch tracking of the deviant stimulus. (A) The autocorrelogram of the response to T1 (top) and T4 (bottom) in the deviant condition for trained participants is plotted here for the pre- and post-training sessions (left and right, respectively). The response autocorrelograms plotted here were derived from the time-domain group average waveforms. (B) The stimulus pitch contour is represented more robustly after training. When collapsing across tones and probability conditions, the average *Frequency Error* dropped for the trained Group, but was more stable in the control. The greatest training-related changes were observed the deviant condition of T1 (top left panel, blue line) ($t(11) = 4.29$, $p = 0.001$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

worse (Session 1 Frequency Error: 43.49 ± 8.15 Hz; Session 2 = 47.99 ± 11.35 Hz, ($t(11) = 1.929$, $p = 0.112$) (Fig. 6). In the case of the control group, the probability effect was not present at retest when collapsing across tones ($F(1, 1) = 0.024$, $p = 0.800$) nor for the individual tones (T1: $t(12) = -0.158$, $p = 0.887$; T4: $t(12) = 0.542$, $p = 0.598$).

4. Discussion

Our results demonstrate that at least two types of neural plasticity are operational in the human auditory brainstem—one that is dependent on the probability with which a sound occurs in the incoming stimulus stream and another that is dependent on the behavioral relevance of the stimulus, which in this case was acquired through short-term auditory training. In the initial test session, when the pitch contours were linguistically irrelevant for all listeners, the stimulus was tracked more *accurately* when it was presented at a higher probability than when it was presented infrequently within the stimulus stream. We interpret this as evidence of probability-dependent plasticity. Yet, after the contour became linguistically meaningful through behavioral training, brainstem representation of the trained feature, the linguistic pitch pattern, became enhanced. This enhancement emerged as better tracking of the pitch pattern after training. We conclude that training-related plasticity, which was more pronounced for the deviant condition compared to the standard one, altered probability-dependent processes in the auditory brainstem.

4.1. Probability-dependent plasticity

We recorded ABRs using an oddball paradigm composed of a standard stimulus that was played repetitively with a high degree of regularity (80%) and a deviant stimulus whose occurrence was both infrequent (20%) and unpredictable. In the initial baseline measurements, we found that stimulus predictability affected the accuracy with which stimulus features were represented in the FFR, with more accurate tracking of the pitch contour observed when the complex pitch pattern occurred frequently vs. infrequently. Previous reports have indicated that FFRs are enhanced to repetitive stimulation (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Skoe & Kraus, 2010b; Slabu et al., 2012) and here we add to the literature by showing that repetition influences not only the amplitude of the response but also the fidelity with which stimulus features are captured in the response. We took a new approach to studying probability-dependent plasticity in the brainstem by using stimuli that were minimally contrastive (same syllable different pitch contour), and where the acoustic contrast carried no lexical meaning (without training). This is in contrast to previous investigations where the stimulus stream was comprised of sounds within the phonemic inventory of the listener (Chandrasekaran et al., 2009; Slabu et al., 2012; Strait, Hornickel, & Kraus, 2011). By adopting this approach, we establish that the brainstem indexes differences in stimulus probability, even when the stimulus contrasts are foreign to the listener.

How do our findings square with the existing literature? Oddball paradigms have been used extensively to study probability-dependent processing in the auditory system in both human and laboratory animals (recently reviewed in: Bendixen, SanMiguel, & Schroger, 2012; Grimm & Escera, 2012). Across the auditory system, repetitive stimulation triggers a phenomenon called stimulus-specific adaptation (SSA), in which neurons reduce their responsivity to a repeating or highly probable stimulus and increase their responsivity to novel stimuli (Bauerle, von der Behrens, Kossl, & Gaese, 2011; Duque, Perez-Gonzalez, Ayala, Palmer, & Malmierca, 2012; Farley, Quirk, Doherty, & Christian, 2010; Malmierca et al., 2009; Perez-Gonzalez et al., 2005; Ulanovsky, Las, Farkas, & Nelken, 2004). SSA has been documented to varying degrees across the auditory system (Nelken & Ulanovsky, 2007), including in the inferior colliculus (IC), a primary neural generator of the FFR (reviewed in: Chandrasekaran & Kraus, 2010). SSA has been hypothesized to be one of the cellular bases of the cortical AEP change-detection component, called the mismatch negativity (MMN) (Grimm & Escera, 2012; Ulanovsky, Las, & Nelken, 2003; Ulanovsky et al., 2004), and it is also thought to underlie enhancements seen in the middle latency response (MLR) to deviant sounds (Grimm, Escera, Slabu, & Costa-Faidella, 2011). Both SSA and MMN are argued to be neural mechanisms for rapidly detecting acoustic changes within the incoming stimulus stream. Rare sounds, especially in the context of repetitively presented sounds, may signify danger, and therefore responding robustly to these deviant sounds may be important for survival. We propose that the brain invokes other mechanisms for sensitizing responses to repetitive stimulation. Repetition-enhancing mechanisms, we theorize, work in concert with SSA and other forms of neural adaptation to capture the statistical signature of the soundscape and the behavioral importance of different auditory objects within that soundscape. Consistent with that idea, Haenschel, Vernon, Dwivedi, Gruzelier, and Baldeweg (2005) describe a "repetition positivity" (RP) that increased progressively as the number of standard stimuli increased (Haenschel et al., 2005). More accurate phase-locking to repetitive stimulation may, we speculate, falls into the same class of repetition-enhancement mechanisms as the RP.

Our findings suggest that the neural generators of the FFR, which include the IC, are sensitive to statistical regularities in the ongoing auditory input. However, unlike what has been shown in single IC units, for the scalp-recorded ABR we found that repetitive stimulation induced greater neural responsivity. Malmierca and colleagues reported that a large fraction of the neurons showing SSA in the IC are onset responders (Malmierca et al., 2009), suggesting that SSA and FFR-enhancements involve different neuronal subtypes. Thus, our findings argue for the independence of SSA and repetition-enhancement mechanisms in the IC. We speculate that both SSA and repetition-enhancements can manifest in the ABR. Based on the response properties of IC neurons, SSA is expected to be more prevalent for the ABR onset response than for the phase-locked FFR component of the ABR. The acoustic characteristics of our stimulus combined with the relatively small number of sweeps (~600/condition) prevented us from effectively examining the ABR onset response. The stimulus /mi/ is composed of a nasal consonant followed by a high-front vowel. Nasal consonants, with their relatively low-energy onsets, do not elicit robust onset responses (Skoe & Kraus, 2010a). However, adaptation of the onset response (in the form of increased latencies and decreased amplitudes) has been reported for other stimuli including trains of clicks and a piano melody (Don, Allen, & Starr, 1977; Lasky, 1997; Skoe & Kraus, 2010b), suggesting that ABR onset reduction may reflect SSA occurring on the single neuron level. However, the findings reported by Skoe and Kraus (2010b) suggest that there may be a "release" from onset adaptation that happens with more prolonged stimulation. Skoe and Kraus (2010b) recorded ABRs to a five-note melody in which the first and second notes were identical. For the first half of the recording (2000 trials, ~45 min), the onset response to the second note was attenuated relative to the first note; however, for the last half of the recording, the onset response amplitude was equated for the first and second notes, suggesting that the initial stage of adaption was overridden by repetition enhancement mechanisms. Interestingly, the FFR to this melody also showed time-dependent enhancements (but no initial adaptation) that grew linearly with increasing number of stimulus presentations. While the Skoe and Kraus (2010b) study did not show evidence for FFR adaptation, other studies suggests that FFR amplitudes do, under certain conditions and in certain populations, adapt to repetitive stimulation, with adaptations being associated with poorer behavioral performance on tests of auditory discrimination (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Skoe et al., 2013). Indeed, Skoe et al. (2013) have recently proposed that there are individual differences in the neural sensitivity to statistical structure that manifests along a continuum from FFR adaptation to FFR enhancement, with individuals who show repetition enhancements being better auditory learners. In the current study, we find that most of the participants fell into the repetition enhancement category.

The exact mechanism by which FFR enhancement occurs is as of yet undetermined. We propose three potential mechanisms to explain how phase-locking was enhanced between the standard and deviant conditions in the baseline recordings. Future experiments are planned to disambiguate these mechanisms, by experimentally varying the number of standard trials separating each deviant. The first proposed mechanism is that phase-locking to the standard stimulus is refined online: through repeated presentation of the standard stimulus phase-locking becomes more accurate resulting in a "clearer" response. By this explanation, there is a "greasing of the wheels" that takes place when a stimulus is presented frequently within a given time window that facilitates the encoding of future presentations of that sound. This explanation would be consistent with the repetition enhancements observed by Haenschel and colleagues (2005) and Skoe and Kraus (2010b), as well as repetition induced increases in cortical-AEP P2 amplitude that occur across successive days (Ross & Tremblay, 2009). But if FFR stimulus-tracking does not become more accurate with increasing number of repetitions, this would suggest a second mechanism, namely that phase-locking is stable to the standard stimulus but is disrupted when a novel sound is encountered. A third possibility is a hybrid of the other two: phase-locking improves as a function of how many stimuli have been presented but the presence of the deviant causes a partial resetting of that process. This third possibility, we predict, would emerge as a slow-building increase in stimulus tracking: within each section of repeating standards, pitch tracking accuracy would build until a deviant sound was encountered. The deviant sound would cause a slight lowering in pitch tracking accuracy for the next standard sound but then pitch tracking accuracy would build again throughout the ensuing train of standards until the next deviant was encountered. This cycle would repeat, leading to an undulating but gradual increase in pitch tracking accuracy over the recording. This type of analysis, however, is challenging given the poor signal to noise ratios of ABRs to individual trials. Techniques such as high-density electrode arrays (i.e., multiple active recording sites) may enable more fine-grained examinations of how the ABR is affected by the statistical properties of the current soundscape and how this sensitivity interacts with prior exposure to soundscapes with different statistical signatures.

4.2. Learning-related plasticity

In addition to probability-related plasticity, we found that the auditory brainstem undergoes a second type of plasticity that is related to sound-to-meaning training. Our results suggest an enhancement of sound patterns that have become behaviorally-relevant through training, consistent with studies from animal models demonstrating brainstem-processing changes following auditory-associative learning (Gao & Suga, 1998; Tzounopoulos & Kraus, 2009) and human models showing that cognitively-based auditory training can drive learning-related plasticity in the developing, maturing, and aging auditory brainstem (Anderson et al., 2013; Carcagno & Plack, 2011; Chandrasekaran et al., 2012; Hornickel, Zecker, Bradlow, & Kraus, 2012; Russo, Nicol, Zecker, Hayes, & Kraus, 2005; Song, Skoe, Banai, & Kraus, 2012; Song et al., 2008).

As part of their training, participants learned an artificial language of 24 novel words. To successfully learn these words, participants had to discern a set of four syllables that were minimally contrasted by their pitch pattern (Fig. 4). In our artificial language, the 8 syllables followed the rules of English phonotactics. Although these 8 syllables are not part of the English lexicon, by including common phonemic combinations, they sound like possible English words. The aspect of the language that was more novel to the native English listeners was the pitch contours. Although English speakers may use syllable-level pitch changes to convey emotion or emphasis, such dynamic pitch changes do not affect the lexical status of a word in English. Through the process of integrating syllable and novel pitch information, the trained participants, we argue, gained implicit knowledge that pitch contours are important semantic cues in this artificial language. By this argument, the pitch contours attained new linguistic status over the course of the training. This change in status may be leading to enhanced pitch tracking (reduced tracking errors) at retest and may be allowing the listeners to generalize to an untrained syllable spoken by an untrained talker. Thus, while the T1 and T4 pitch contours do not carry lexical meaning perse (i.e., the pitch contour alone is not sufficient for distinguishing the 24 words in the artificial language), we maintain that the contour became linguistically meaningful to the listeners. On an average, participants showed a consistent increase in word learning from session 1 to session 5. Between sessions 5 and 9, learning stabilized and participants reached a plateau, suggesting that the pitch patterns became and stayed linguistically relevant for the participants. In addition, because the pitch contours were learned as part of a behavioral training paradigm, we further argue that the T1 and T4 pitch contours became behaviorally meaningful well.

In the control group, however, pitch tracking did not improve after training. This outcome is consistent to what has been reported previously (Carcagno & Plack, 2011; Song, Nicol, & Kraus, 2011; Song et al., 2012). However, given that the control group was not provided any additional auditory stimulation between sessions 1 and 2, we leave open the possibility that the observed effects in the trained group resulted from greater exposure to the pitch contours. Thus, increased pitch tracking may reflect the formation of an implicit memory trace but not necessarily reflect explicit learning. By not including a control condition in which the exposure to the stimuli matched the trained group, this tempers how we interpret our results. Future investigations should also examine the specificity of training by including an active control group that is trained in a similar fashion but on a different, unrelated auditory feature (i.e., a feature other than the pitch contour) and/or by including an active control group that undergoes a different type of training that is matched in time and computer use to the experimental group. In a recent training study that included such an active control group, Anderson et al. (2013) showed that 8-weeks of in-home computerized auditory training led to systematic decreases in ABR timing in an older-adult population. In contrast, the active control group who participated in an in-home education program for the same time period did not demonstrate ABR plasticity (Anderson et al., 2013). While there are methodological differences that complicate generalizations between our study and the Anderson et al., 2013 one, we view the Anderson et al. study as robust evidence that training-related ABR plasticity can result from training and not other unspecified factors.

4.3. Interaction between probability-dependent and training-related brainstem plasticity: hierarchy of experience

We found an interaction between probability-dependent and training-related plasticity. Notably, training had the greatest influence on the pitch contour that was most poorly encoded at pretest (T1), especially when that sound (T1) was presented as the deviant. Thus, given that deviant sounds were represented less robustly than the standard tones at pretest, we conclude that the effect of training was highly specific and that it strengthened what might be considered a physiological weakness. Another explanation for why the training effect was specific to T1-deviants and why it was not observed for T4-deviants is that T4, when presented in the context of repeating T1-standards, is a less salient acoustic contrast than the reversed condition, given that the starting pitch of T4 is almost equivalent to the ending pitch of T1 Fig. 2. In contrast, T1 in the context of T4 is more salient because there is a bigger change in pitch between stimuli, making T1 easier to detect. It is, therefore, possible that training gave rise to enhanced change detection for the contrast that was more acoustically salient.

Neural sensitivity to stimulus-probability is thought to reflect on-line processes that monitor the sensory input for changes in the auditory environment. The ability to detect auditory change in a near instantaneous fashion is essential for survival (Naatanen, Paavilainen, Rinne, & Alho, 2007), especially given that the auditory world is not stationary. Training-related brainstem plasticity, on the other hand, may relate to improved neural representation of sounds that have acquired importance through repeated exposure and behavioral reinforcement (Gilbert, Li, & Piech, 2009; Gilbert & Sigman, 2007; Suga, 2008). While probability-dependent and training-related plasticity likely reflect distinct biological needs, our findings suggest that they do not operate independently. Indeed, we show that both affect how the pitch contour of the stimulus is represented. Based on the outcomes of the baseline measurements, we conclude that infrequent sounds are encoded more weakly than frequent ones, but that through training, pitch tracking of the infrequent sounds can be improved. In the trained group, this improvement resulted in deviant sounds being processed more accurately than common sounds. In essence, the effect of probability was reversed for T1 after training.

Our findings lead us to conclude that the sound-to-meaning training altered on-line probability calculations, leading to a change in how the novel sound within the oddball sequence was processed. Across this experiment, two layers of novelty were manipulated. By using an oddball paradigm, in which one sound was played more often than the other, we manipulated the novelty of a sound within a short time window (roughly 20 min). We also manipulated the overall novelty of the stimulus by using non-native pitch contours to which all listeners were initially naïve. Then, half of the participants learned to distinguish the two pitch contours within a linguistic context. Thus, after training, the pitch contours were at the same time less novel and more linguistically-meaningful to that subset of participants. As a consequence, the training process may have skewed how the oddball sequence was processed because deviant sounds, although played with a low probability within the immediate sensory experience (20%), had become less novel sounding to the trained listeners through recent past experience. For the control group, the stimuli were also presumably less foreign sounding at retest. This increased familiarity may explain why probability effects were not found when the control group was retested.

We propose that the neural mechanisms associated with on-line probability detection can be changed by short-term experience. We invoke this proposition not only to explain our own dataset but to also explain the asymmetrical effect reported by Slabu and colleagues (Slabu et al., 2012). Slabu et al. demonstrated that the human auditory brainstem can detect infrequent sounds within a train of repetitive stimuli, as reflected by larger FFRs to standard vs. deviant conditions. However, the effect did not generalize across stimuli. They used an oddball sequence composed of two speech sounds that were part of the listener's native language, but where the phonemic category of one of the sounds was more ambiguous. Probability-dependent plasticity was not found for this ambiguous stimulus, suggesting that categorical perception, which is an experience-dependent process (Kuhl, Tsao, & Liu, 2003), can affect probability calculations (Bidelman, Hutka, & Moreno, 2013). Based on the small collection of research on probability-dependent brainstem plasticity (Chandrasekaran et al., 2009; Parbery-Clark et al., 2011; Skoe et al., 2013; Slabu et al., 2012; Strait et al., 2011), we postulate that auditory experience serves to either deemphasize differences or maximize differences between frequent and infrequent sounds, with the extent and direction of the change being dictated by the nature, duration, and recency of the experience.

The notion of experience-dependent brainstem processing is not new. Provocative examples can be found in both the animal and human literatures (de Boer & Thornton, 2008; Kraus & Chandrasekaran, 2010; Krishnan et al., 2005; Krizman, Marian, Shook, Skoe, & Kraus, 2012). However, little is known about how different kinds of experience may combine to shape brainstem processing. We theorize that experience acts in a hierarchical fashion, with some experiences being subordinate to others. Within the context of our experiment, it would appear that recent past experience trumps the immediate sensory context. Two alternative (but not necessarily mutually exclusive) explanations might be that more extensive experience trumps more transient experience and/or that active listening experience trumps more passive listening experiences. In other words, 9 sessions of explicit auditory training trumps the more transient experience of passively listening to the oddball conditions during electrophysiological testing. The next generation of ABR-related research should test this hierarchy of experience theory more directly by layering different kinds of experience within a single experiment. For example, by administering our experimental paradigm to auditory experts such as musicians or administering two types of auditory training back-to-back, it may be possible to which types of experiences are additive and which are competitive.

4.4. Neural origins and mechanisms: The interaction between probability-dependent and learning-dependent plasticity

Although the exact neural mechanisms underlying brainstem plasticity are difficult to directly discern using non-invasive recordings in humans, based on what is known from the animal literature and current theories of human brainstem plasticity (Kraus & Chandrasekaran, 2010; Krishnan & Gandour, 2009; Tzounopoulos & Kraus, 2009), we can offer some potential insight into the mechanisms that lead to the interaction between probability-dependent and learning-dependent plasticity in the auditory brainstream.

The IC, one of the most metabolically active centers of the brain (Sokoloff, 1977), serves as a relay center for ascending and descending auditory information (Winer, 2006). The extensive descending pathways, known as the corticofugal pathways, are the conduit by which the cortex can send instructive feedback to lower auditory regions leading to an enhancement or suppression of subcortical activity (Bajo & King, 2012; Winer, 2006). Thus, one explanation for our findings is that probability-dependent plasticity requires input from the auditory cortex (Bauerle et al., 2011; Nelken & Ulanovsky, 2007) and that auditory training has changed the instructive feedback that the cortex sends to the brainstem. By this explanation, the auditory cortex dictates how the brainstem responds to infrequent sounds, both before and after training. An alternative explanation is that top-down feedback is not necessary for probability-dependent plasticity (Antunes & Malmierca, 2011; Malmierca et al., 2009) but that learning-related top-down feedback can override the local brainstem mechanisms that subserve probability detection. If this explanation holds, then short-term training may have led to a shift from a more bottom-up approach to a more top-down approach to probability detection. While speculative at this point, we theorize that experiences that lead to a shift in the balance between bottom-up and top-down mechanisms, and/or that alter the instructive feedback from the cortex, have a higher position in the hierarchy of experience than those that maintain the status quo.

5. Conclusion

Our results demonstrate the interaction between two forms of plasticity in the mature human auditory brainstem—an immediate, within-session plasticity related to stimulus-probability, and a training-related plasticity that is driven by increased signal relevance. We show that increased signal relevance results in the deviant stimulus being processed as if it were less deviant.

Acknowledgments

This work was supported by National Science Foundation Grant BCS-0719666 and National Institute on Deafness and Other Communication Disorders Grant R01 DC008333. We thank Trent Nicol, Adam Tierney, and Jennifer Krizman for their feedback on a previous version of this manuscript.

References

Akhoun, I., Gallego, S., Moulin, A., Menard, M., Veuillet, E., Berger-Vachon, C., et al. (2008a). The temporal relationship between speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/ in normal-hearing adults. *Clinical Neurophysiology, 119*, 922–933.

Akhoun, I., Moulin, A., Jeanvoine, A., Menard, M., Buret, F., Vollaire, C., et al. (2008b). Speech auditory brainstem response (speech ABR) characteristics depending on recording conditions, and hearing status: An experimental parametric study. *Journal of Neuroscience Methods, 175*, 196–205.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 4357–4362.

Antunes, F. M., & Malmierca, M. S. (2011). Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body. *Journal of Neuroscience, 31*, 17306–17316.

Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits, 6*, 114.

Bauerle, P., von der Behrens, W., Kossl, M., & Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *Journal of Neuroscience, 31*, 9708–9722.

Bendixen, A., SanMiguel, I., & Schroger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology, 83*, 120–131.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*, 425–434.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS One, 8*, e60676.

Carcagno, S., & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology, 12*, 89–100.

Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron, 64*, 311–319.

Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology, 47*, 236–246.

Chandrasekaran, B., Kraus, N., & Wong, P. C. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of Neurophysiology, 107*, 1325–1336.

de Boer, J., & Thornton, A. R. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience, 28*, 4929–4937.

Don, M., Allen, A. R., & Starr, A. (1977). Effect of click rate on the latency of auditory brain stem responses in humans. *Annals of Otology, Rhinology, and Laryngology, 86*, 186–195.

Duque, D., Perez-Gonzalez, D., Ayala, Y. A., Palmer, A. R., & Malmierca, M. S. (2012). Topographic distribution, frequency, and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *Journal of Neuroscience, 32*, 17762–17774.

Farley, B. J., Quirk, M. C., Doherty, J. J., & Christian, E. P. (2010). Stimulus-specific adaptation in auditory cortex is an NMDA-independent process distinct from the sensory novelty encoded by the mismatch negativity. *Journal of Neuroscience, 30*, 16475–16484.

Gao, E., & Suga, N. (1998). Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 12663–12670.

Gao, E., & Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: Role of the corticofugal system. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 8081–8086.

Gilbert, C. D., Li, W., & Piech, V. (2009). Perceptual learning and adult cortical plasticity. *Journal of Physiology, 587*, 2743–2751.

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron, 54*, 677–696.

Grimm, S., & Escera, C. (2012). Auditory deviance detection revisited: Evidence for a hierarchical novelty system. *International Journal of Psychophysiology, 85*, 88–92.

Grimm, S., Escera, C., Slabu, L., & Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology, 48*, 377–384.

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience, 25*, 10494–10501.

Herholz, S. C., Boh, B., & Pantev, C. (2011). Musical training modulates encoding of higher-order regularities in the auditory cortex. *European Journal of Neuroscience, 34*, 524–529.

Hornickel, J., Zecker, S. G., Bradlow, A. R., & Kraus, N. (2012). Assistive listening devices drive neuroplasticity in children with dyslexia. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 16731–16736.

Jacobsen, T., & Schroger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology, 114*, 1133–1143.

Jacobsen, T., Schroger, E., Horenkamp, T., & Winkler, I. (2003). Mismatch negativity to pitch change: Varied stimulus proportions in controlling effects of neural refractoriness on human auditory event-related brain potentials. *Neuroscience Letters, 344*, 79–82.

Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schroger, E., & Jacobsen, T. (2008). Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *International Journal of Psychophysiology, 67*, 131–143.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*, 599–605.

Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language, 110*, 135–148.

Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: Representation of pitch contours in Chinese tones. *Hearing Research, 189*, 1–12.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research Cognitive Brain Research, 25*, 161–168.

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 7877–7881.

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 9096–9101.

Kujala, T., & Naatanen, R. (2010). The adaptive brain: A neurophysiological perspective. *Progress in Neurobiology, 91*, 55–67.

Lasky, R. E. (1997). Rate and adaptation effects on the auditory evoked brainstem response in human newborns and adults. *Hearing Research, 111*, 165–176.

Malmierca, M. S., Cristaudo, S., Perez-Gonzalez, D., & Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *Journal of Neuroscience, 29*, 5483–5493.

Malone, B. J., & Semple, M. N. (2001). Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *Journal of Neurophysiology, 86*, 1113–1130.

Naatanen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing, 16*, 6–18.

Naatanen, R. (2008). Mismatch negativity (MMN) as an index of central auditory system plasticity. *International Journal of Audiology, 47*(Suppl. 2), S16–20.

Naatanen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*, 2544–2590.

Naatanen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). ''Primitive intelligence'' in the auditory cortex. *Trends in Neuroscience, 24*, 283–288.

Nelken, I., & Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology, 21*, 214–223.

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia, 49*, 3338–3345.

Perez-Gonzalez, D., Malmierca, M. S., & Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience, 22*, 2879–2885.

Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research, 248*, 48–59.

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioral Brain Research, 156*, 95–103.

Skoe, E., & Kraus, N. (2010a). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing, 31*, 302–324.

Skoe, E., & Kraus, N. (2010b). Hearing it again and again: On-line subcortical plasticity in humans. *PLoS One, 5*, e13645.

Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience, 243*, 104–114.

Slabu, L., Escera, C., Grimm, S., & Costa-Faidella, J. (2010). Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials. *European Journal of Neuroscience, 32*, 859–865.

Slabu, L., Grimm, S., & Escera, C. (2012). Novelty detection in the human auditory brainstem. *Journal of Neuroscience, 32*, 1447–1452.

Sokoloff, L. (1977). Relation between physiological function and energy metabolism in the central nervous system. *Journal of Neurochemistry, 29*, 13–26.

Song, J. H., Nicol, T., & Kraus, N. (2011). Test-retest reliability of the speech-evoked auditory brainstem response. *Clin. Neurophysiol., 122*, 346–355.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex, 22*, 1180–1190.

Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience, 20*, 1892–1902.

Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions, 7*, 44.

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A, 194*, 169–183.

Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: Recent progress. *Journal of the National Academy of Sciences of the United States of America, 97*, 11807–11814.

Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron, 36*, 9–18.

Szymanski, F. D., Garcia-Lazaro, J. A., & Schnupp, J. W. (2009). Current source density profiles of stimulus-specific adaptation in rat auditory cortex. *Journal of Neurophysiology, 102*, 1483–1490.

Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron, 62*, 463–469.

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience, 24*, 10440–10453.

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience, 6*, 391–398.

Winer, J. A. (2006). Decoding the auditory corticofugal systems. *Hearing Research, 212*, 1–8.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.