J Neurophysiol 117: 1409–1424, 2017.
First published January 18, 2017; doi:10.1152/jn.00445.2016.

RESEARCH ARTICLE | *Auditory System Plasticity*

Stability and plasticity in neural encoding of linguistically relevant pitch patterns

Zilong Xie,[1]* Rachel Reetzke,[1]* and Bharath Chandrasekaran[1,2,3,4,5]

[1]*Department of Communication Sciences and Disorders, The University of Texas at Austin, Austin, Texas;* [2]*Department of Psychology, The University of Texas at Austin, Austin, Texas;* [3]*Department of Linguistics, The University of Texas at Austin, Austin, Texas;* [4]*Institute for Neuroscience, The University of Texas at Austin, Austin, Texas; and* [5]*Institute for Mental Health Research, The University of Texas at Austin, Austin, Texas*

Submitted 6 June 2016; accepted in final form 9 January 2017

Xie Z, Reetzke R, Chandrasekaran B. Stability and plasticity in neural encoding of linguistically relevant pitch patterns. *J Neurophysiol* 117: 1409–1424, 2017. First published January 18, 2017; doi:10.1152/jn.00445.2016.—While lifelong language experience modulates subcortical encoding of pitch patterns, there is emerging evidence that short-term training introduced in adulthood also shapes subcortical pitch encoding. Here we use a cross-language design to examine the stability of language experience-dependent subcortical plasticity over multiple days. We then examine the extent to which behavioral relevance induced by sound-to-category training leads to plastic changes in subcortical pitch encoding in adulthood relative to adolescence, a period of ongoing maturation of subcortical and cortical auditory processing. Frequency-following responses (FFRs), which reflect phase-locked activity from subcortical neural ensembles, were elicited while participants passively listened to pitch patterns reflective of Mandarin tones. In *experiment 1*, FFRs were recorded across three consecutive days from native Chinese-speaking ($n = 10$) and English-speaking ($n = 10$) adults. In *experiment 2*, FFRs were recorded from native English-speaking adolescents ($n = 20$) and adults ($n = 15$) before, during, and immediately after a session of sound-to-category training, as well as a day after training ceased. *Experiment 1* demonstrated the stability of language experience-dependent subcortical plasticity in pitch encoding across multiple days of passive exposure to linguistic pitch patterns. In contrast, *experiment 2* revealed an enhancement in subcortical pitch encoding that emerged a day after the sound-to-category training, with some developmental differences observed. Taken together, these findings suggest that behavioral relevance is a critical component for the observation of plasticity in the subcortical encoding of pitch.

NEW & NOTEWORTHY We examine the timescale of experience-dependent auditory plasticity to linguistically relevant pitch patterns. We find extreme stability in lifelong experience-dependent plasticity. We further demonstrate that subcortical function in adolescents and adults is modulated by a single session of sound-to-category training. Our results suggest that behavioral relevance is a necessary ingredient for neural changes in pitch encoding to be observed throughout human development. These findings contribute to the neurophysiological understanding of long- and short-term experience-dependent modulation of pitch.

pitch encoding; experience-dependent plasticity; auditory training; adolescence; frequency-following response

PITCH IS A CRITICAL COMPONENT for the perception of prosody in speech and serves as a salient cue for speech segregation. For speakers of tone languages, change in pitch patterns within a syllable can change word meaning. The neural encoding of linguistically relevant pitch patterns can be assessed using the scalp-recorded frequency-following response (FFR), a preattentive electrophysiological metric that captures phase-locked responses from subcortical ensembles along the auditory pathway (Chandrasekaran and Kraus 2010; Coffey et al. 2016; Smith et al. 1975; Sohmer et al. 1977). The extent of similarity between the pitch pattern extracted from the stimulus and the FFR reflects the fidelity of pitch representation in the human auditory system (see, e.g., Chandrasekaran et al. 2012; Coffey et al. 2016; Krishnan et al. 2004; Krishnan et al. 2005; Wong et al. 2007).

Previous studies have examined FFRs in native speakers of tone languages (e.g., Mandarin Chinese speakers; Bidelman et al. 2011; Jeng et al. 2011b; Krishnan et al. 2005; Krishnan et al. 2009a; Krishnan et al. 2009b; Krishnan et al. 2010b; Xu et al. 2006) relative to nonnative speakers (e.g., English speakers) to investigate the extent to which long-term language experience shapes neural encoding of pitch. Collectively, these studies have revealed that the fidelity of neural tracking of dynamic pitch patterns is more robust in native tone language speakers. This evidence suggests that individuals' long-term language experience modulates preattentive, subcortical encoding of pitch (Kraus et al. 2009). Short-term auditory training can also modify subcortical representations of pitch patterns, even in adulthood (Carcagno and Plack 2011; Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008). These studies have provided evidence that subcortical plasticity is not limited to a critical developmental period but rather is retained throughout the life span. Using the FFR as a metric, this body of literature has rewritten the conceptual understanding of human auditory plasticity beyond the so-called "critical period" of auditory development.

The focus of the present study is to address three critical, unanswered questions related to experience-dependent audi-

---

* Z. Xie and R. Reetzke contributed equally to this work.

Address for reprint requests and other correspondence: B. Chandrasekaran, Univ. of Texas at Austin, 2504A Whitis Ave. (A1100), Austin, TX 78712 (e-mail: bchandra@utexas.edu).

tory plasticity. In *experiment 1* we ask: How stable is the plasticity attributed to long-term language experience? Recent investigations using the FFR as a metric have reported plastic changes in sensory encoding within a session of recording (Skoe et al. 2013, 2015b). Although prior work that has examined the FFR has yielded robust test-retest reliability (Hornickel et al. 2012; Russo et al. 2004; Song et al. 2011), the stability of long-term language-related plasticity over multiple days of recording has not been evaluated. In *experiment 2* we ask: What is the timescale of short-term training-related plasticity? Previous studies that have examined the FFR to pitch patterns have revealed experience-dependent plasticity of subcortical encoding that followed several days to weeks of extensive training (Carcagno and Plack 2011; Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008). While these studies have established that subcortical plasticity persists in the mature auditory system, the extent to which a single day of sound-to-category training modulates subcortical encoding of linguistically relevant pitch patterns remains unclear. There is an additional paucity of evidence for auditory plasticity and related behavioral outcomes in adolescence, a period of human development characterized by heightened changes in neural structure and function (Davidow et al. 2016; Luna et al. 2010; Steinberg 2008).

With respect to auditory development, cross-sectional and longitudinal studies have revealed changes in electrophysiological responses to sound throughout adolescence, providing evidence for continued maturation of subcortical and cortical auditory processing (Albrecht et al. 2000; Bishop et al. 2007; Fitzroy et al. 2015; Krizman et al. 2015; Mahajan and McArthur 2012; Ponton et al. 2000; Skoe et al. 2015a). With respect to behavior, classic interpretations of a "critical period" for language acquisition posit that second language learning in adolescents and adults will be slower and less successful relative to first language learning, because of neural commitment to the first language. (Krashen 1973; Lenneberg et al. 1967; Snow and Hoefnagel-Höhle 1978). Within this interpretation, not much learning distinction is expected between postpubescent adolescent and adult acquisition of novel speech categories. However, recent evidence has revealed that heightened sensitivity of reward processing mediated by striatal learning systems in adolescence confers feedback-based learning advantages relative to adults (Davidow et al. 2016; DiMenichi and Tricomi 2016). The extent to which these feedback-based learning advantages during adolescence generalize to advantages in novel speech category learning has yet to be tested. Therefore, we ask a third question: Are there training-related differences in subcortical pitch processing and behavioral learning gains between adolescents and adults? To track the dynamics of auditory plasticity as a function of feedback-based sound-to-category training in adolescents and adults, we measured FFRs before, during, and immediately after a single session of training, as well as a day after training. We motivate each of these experimental pursuits below.

Experiment 1 examined the stability of language experience-dependent neural plasticity to linguistically relevant pitch patterns. A series of cross-language studies that have compared Chinese and English listeners' FFRs to Mandarin tones (an example of linguistically relevant pitch patterns) have shown that Chinese listeners neurally track dynamic changes in pitch with greater precision relative to native English listeners (for review see Krishnan and Gandour 2009). Dynamic pitch changes within syllables are more prevalent in tonal languages like Mandarin Chinese relative to nontonal languages like English (Gandour 1983; Gandour and Harshman 1978). Cross-language differences in the neural encoding of linguistically relevant pitch patterns have been attributed to local reorganization of subcortical auditory circuitry that results in more robust encoding of frequently occurring signal parameters (e.g., dynamic pitch patterns) during language development (Bidelman et al. 2011; Krishnan et al. 2005; Krishnan et al. 2009a; Krishnan et al. 2012; Krishnan and Gandour 2009; Xu et al. 2006). A large body of animal studies have demonstrated that local reorganization during development persists through adulthood (Keuroghlian and Knudsen 2007; Linkenhoker et al. 2005). Therefore, per the local reorganization account, there is considerable stability in long-term language experience-dependent neuroplasticity. This stability is a direct result of hardwired changes in circuitry within the subcortical auditory system.

Additionally, it has been demonstrated that short-term passive exposure may also lead to reorganization local to the subcortical auditory circuitry. For example, in animal models subcortical plasticity to stimulus regularities can occur with remarkably few instances (Dahmen et al. 2010; Dean et al. 2005; Robinson et al. 2016). Animal studies have also revealed that repeated sound exposure can lead to changes in the function of subcortical neurons, specifically at the level of the auditory midbrain (see, e.g., Yan and Suga 1998). In humans, passive exposure has been found to induce plastic changes to key signal features in early visual cortices (i.e., V1) (Watanabe et al. 2002a, 2002b). These visual regions can be likened to the auditory midbrain (i.e., inferior colliculus) with respect to the nature and complexity of sensory processing (Nelken 2004). These findings point to a possibility of plastic changes in the subcortical encoding of linguistically relevant pitch patterns following passive sound exposure.

To test the stability of lifelong language experience-dependent plasticity, we recorded FFRs to linguistic pitch patterns from Chinese and English listeners across 3 days of repeated recording sessions. Ten native speakers of Chinese and ten native speakers of English completed three recording sessions each. Three different hypotheses based on prior work were generated for *experiment 1*: *1*) Chinese listeners would show enhanced FFRs relative to English listeners, which would remain stable through multiple days of recording; *2*) Chinese listeners would have a robust FFR advantage in *day 1* that would reduce over 3 days of recording because of enhanced plasticity in English listeners; or *3*) there would be a main effect of linguistically relevant pitch pattern exposure, such that there would be more robust FFRs for both participant groups on *day 3* relative to *day 1*. To anticipate, we found robust and stable cross-language differences (Chinese > English) in the fidelity of FFRs to linguistic pitch patterns. We found no evidence of plasticity to linguistic pitch patterns induced by passive exposure in both native Chinese and English speakers. These results speak to the stability of lifelong language experience-dependent plasticity.

Experiment 2 examined the impact of short-term training on experience-dependent neural plasticity to linguistically relevant pitch patterns. In *experiment 1*, we assessed long-term language experience-dependent plasticity for the neural representation of linguistically relevant pitch patterns. We found remarkable stability in FFR metrics over the timescale of multiple days for both participant groups. Even after native English speakers (for whom linguistic pitch patterns were not behaviorally relevant) listened to several thousand trials of repetitive presentation of the stimuli, no change was observed in the neural representation of the pitch patterns. In *experiment 2*, we aimed to establish behavioral relevance for the linguistic pitch patterns with a single session of sound-to-category training in native English-speaking adolescents and adults. We then examined the extent to which the induced behavioral relevance led to changes in neural encoding of linguistic pitch patterns, as reflected by the FFR collected before, during, and immediately after training, as well as a day after training ceased.

For *experiment 2*, our aims were twofold. First, we aimed to investigate the extent to which behavioral relevance induced by a short period of active sound-to-category training leads to plastic changes in subcortical pitch encoding. By recording FFRs before, during, and after training, we attempted to identify the time point for the emergence of subcortical plasticity. Second, we examined the extent to which adolescents and adults differ in: *1*) learning novel linguistically relevant speech patterns with feedback-based sound-to-category training and *2*) subcortical plasticity related to the sound-to-category training. Below we review prior work that has examined the dynamics of subcortical auditory plasticity and specify how *experiment 2* extends previous work on human auditory plasticity.

Studies on mature animal models have shown that the corticofugal system, originating from the auditory cortex, forms multiple feedback loops from the cortex to subcortical structures (for review see Suga 2008, 2012; Suga et al. 2002). These cortical-to-subcortical feedback loops play a critical role in the modulation of subcortical auditory plasticity to behaviorally relevant signals (for review see Suga 2008; Suga 2012; Suga et al. 2002). Although plasticity can occur at a single neuronal level as a function of auditory repetition (Malmierca et al. 2009; Yan and Suga 1998), plastic changes are more substantial and specific when the auditory stimulation is behaviorally relevant (Yan and Suga 1998). Enhanced plasticity as a function of behavioral relevance has been referred to as "egocentric tuning" and is reflected by enhanced representation of signal features that are deemed to be of critical importance to the animal (Suga 2008; Suga 2012; Suga et al. 2002; Yan and Suga 1996; Zhang et al. 1997). How are these behaviorally relevant cortical constructs established in the adult animal? Based on a plethora of studies on auditory deprivation and congenital hearing loss (Kral et al. 2000; Kral et al. 2002; Kral et al. 2005), Kral and Eggermont (2007) posit that during development passive bottom-up mechanisms are highly labile while cortical structures are still undergoing maturation. In turn, behaviorally relevant representations in cortical areas emerge through the action of subcortical-to-cortical circuitry throughout development. Consequently, once the cortical structures reach maturity, top-down corticofugal control of subcortical structures may be more dominant in the modulation of subcortical auditory processing. Based on this body of animal work, it has been suggested that subcortical processing of speech reflected by the FFR in humans may be influenced by the maturity of the corticofugal pathway (Banai et al. 2009; Chandrasekaran et al. 2014b; Johnson et al. 2008).

In humans, structural data show that maturation of corticofugal and cortical-cortical connectivity is ongoing during adolescence (Paus et al. 1999). Functionally, cross-sectional evidence has demonstrated that early childhood (i.e., 5–11 yr) represents a period of heightened subcortical auditory processing relative to adulthood, where the FFR is larger in amplitude, earlier in latency, and more consistent than young adult FFRs (Krizman et al. 2015; Skoe et al. 2015a). Based on the assumption that plasticity of auditory subcortical processing declines beyond childhood and cerebral lateralization of language function occurs at puberty (Johnson and Newport 1989; Johnson and De Haan 2015; Lenneberg et al. 1967; Pujol et al. 1999), it has been argued that second-language learning must begin early in life for a child to learn a new language (see, e.g., Asher and García 1969; Hakuta et al. 2003). Contrary to this supposition, recent evidence has demonstrated that auditory training (3 yr of in-school music training) leads to better language skills thought to arise from enhanced subcortical speech processing (Tierney et al. 2015). Prior studies have also demonstrated that training-related plasticity in the subcortical neural representation of linguistically relevant pitch patterns can occur even in adulthood (Carcagno and Plack 2011; Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008). For example, after more than a week of training on a sound-to-meaning paradigm, enhanced pitch tracking was observed for dynamic linguistic pitch patterns (Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008).

The studies reviewed provide evidence that neural plasticity is not hindered by neural maturation and is indeed possible after different forms and lengths of auditory training. However, several questions regarding the nature of auditory plasticity remain unanswered. First, the time of training needed for neural and behavioral changes to occur is unclear. Do plastic changes require extensive training over multiple days, weeks, or years? Is a period of learning consolidation necessary to observe neural changes (Alain et al. 2015)? Additionally, no study to date has examined how developmental differences in auditory processing observed between adolescence and adulthood relate to the ability to learn novel speech categories, and consequently no study has examined the extent to which speech sound-to-category training induces subcortical plasticity differences in adolescents relative to adults. Are adolescents able to capitalize on auditory experience-dependent plasticity more than adults because of inherent developmental malleability? Or is subcortical plasticity greater in adults relative to adolescents because of more mature corticofugal connectivity? In *experiment 2*, we sought to address all of these questions. We elicited FFRs from native English-speaking adolescents and adults before, during, and immediately after a session of sound-to-category training. Crucially, we probed the retention of training-related changes by assessing FFRs and behavioral gains a day after training ceased.

To induce behavioral relevance for the linguistic pitch patterns in the native English speaking participants, we implemented a sound-to-category training paradigm that has been extensively studied with neuroimaging (Yi et al. 2016) as well as computational modeling methods (Chandrasekaran et al. 2014a; Chandrasekaran et al. 2015; Chandrasekaran et al. 2016; Maddox and Chandrasekaran 2014; Maddox et al. 2013; Smayda et al. 2015). These approaches have demonstrated significant changes in neural processing (Yi et al. 2016) and reductions in perceptual variability (Smayda et al. 2015) within just a few hundred trials of feedback-based training in a single day. Within a session of sound-to-category training, neural circuitry underlying reward processing is more heavily engaged, which indicates a change in the behavioral relevance of the trained stimuli (Yi et al. 2016). Furthermore, computational modeling of the behavioral results has shown that just a single session of sound-to-category training can lead to an increase in perceptual selectivity (corresponding to a reduction in perceptual noise) to linguistic pitch patterns (Smayda et al. 2015). These studies point to the likelihood that a single session of sound-to-category training leads to neural plasticity in sensory processing of dynamic pitch patterns.

To summarize, in *experiment 1* we aimed to examine the stability of long-term language experience-dependent plasticity for the encoding of linguistic pitch patterns. In *experiment 2* we examined the extent to which induced behavioral relevance led to neural plasticity in the encoding of linguistic pitch patterns before, during, and immediately after a single session of sound-to-category training, as well as a day after training ceased. We further investigated the extent to which adolescents and adults differed on learning novel linguistically relevant speech categories and subcortical plasticity associated with learning gains. For *experiment 1*, we hypothesized that long-term language experience-dependent plasticity (Chinese > English) would be stable across multiple days of recording, based on the local reorganization theory (Keuroghlian and Knudsen 2007; Knudsen 2002; 1999). For *experiment 2*, we hypothesized that: *1*) in line with animal studies (Yan and Suga 1998), once behavioral relevance was established for the linguistic pitch patterns subcortical plasticity, as reflected by the FFR to the linguistic pitch patterns, would be observed for both adolescents and adults; *2*) adolescents would demonstrate better feedback-based sound-to-category learning due to heightened sensitivity to reward processing (Davidow et al. 2016) and *3*) behavioral learning gains would be associated with more robust subcortical encoding of the linguistically relevant stimuli in adolescents relative to adults.

MATERIALS AND METHODS

General Materials and Methods

The following materials and methods are common to *experiments 1* and *2*. Written informed consent was obtained from all adult participants, as well as from adolescents and parents of adolescents. All materials and procedures were approved by the Institutional Review Board at the University of Texas at Austin. All procedures were carried out in accordance with approved guidelines. All participants received $15/h monetary compensation for their participation.

Electrophysiological testing. STIMULI. We collected FFRs to Mandarin tones: *tone 1* and *tone 2* (Fig. 1*A*). A native male speaker of Mandarin Chinese from Beijing produced both tones in the context of the same syllable /a/. The tones were recorded at a sampling rate of 44.1 kHz. The two tones differed in fundamental frequency (F0) contours: *tone 1* has a relatively flat F0 contour, and *tone 2*, in comparison, has a rising F0 contour (Fig. 1*A*). The two tones were duration normalized at 270.8 ms.

EEG RECORDING. In each recording session, participants sat in an acoustically shielded booth and watched a muted movie or television show of their choice with subtitles. The stimuli were binaurally presented in alternating polarity via insert earphones (ER-3; Etymotic Research, Elk Grove Village, IL) at ~80 dB sound pressure level (SPL). Stimuli were presented with a variable interstimulus interval of 156–673 ms. The two tones were presented in separate blocks, and the order of blocks was counterbalanced across participants. The stimulus presentation was controlled by E-Prime 2.0.10 software (Schneider et al. 2002). Participants were instructed to ignore the sounds, focus on the selected movie or television show, and refrain from extraneous movement. Electrophysiological responses to *tones 1* and *2* were collected from the scalp (Cz) using Brain Vision PyCorder

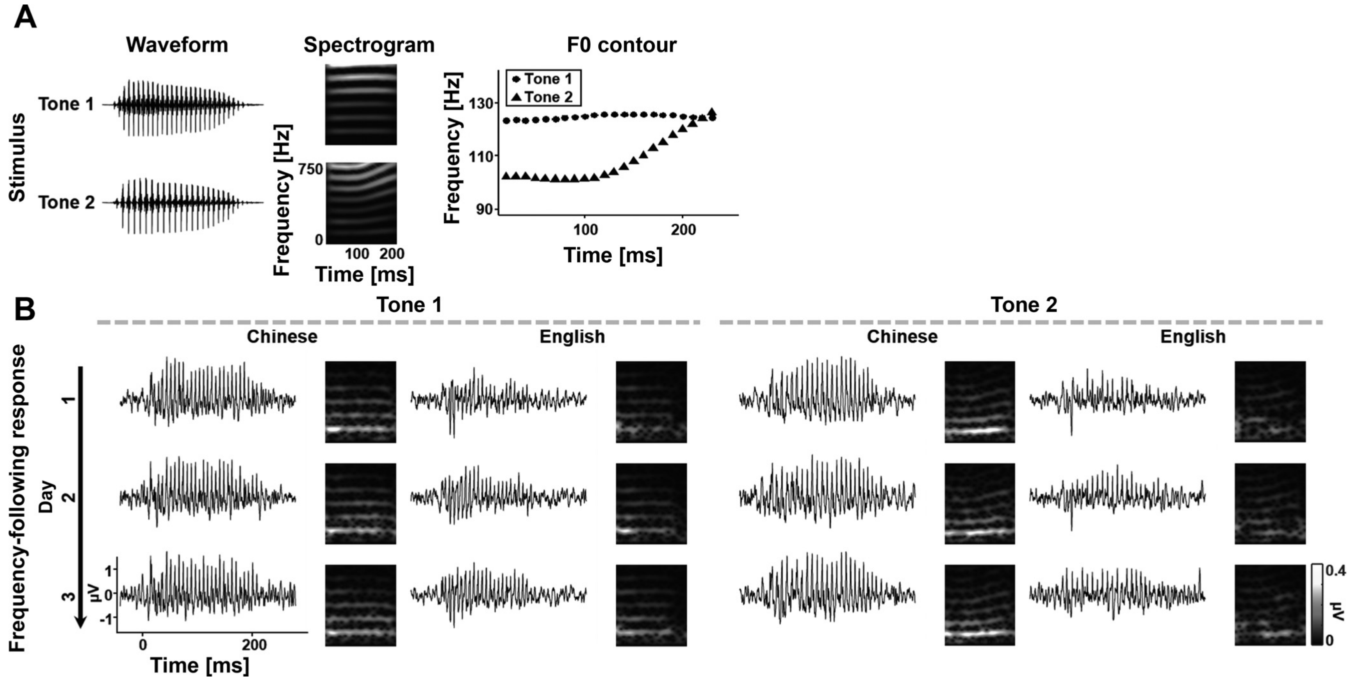

Fig. 1. *A*: waveforms, spectrograms, and fundamental frequency (F0) contours of the stimuli: a flat (*tone 1*) and a rising (*tone 2*) linguistically relevant pitch pattern, used to elicit the frequency-following response (FFR). *B*: waveforms and spectrograms of FFRs elicited by *tone 1* and *tone 2* across 3 days in an example native Chinese participant and an example native English participant in *experiment 1*. Note that the FFRs are highly stable across days within participants.

1.0.7 (Brain Products, Gilching, Germany) with Ag-AgCl scalp electrodes, with the right mastoid as reference, and the left mastoid as ground. Responses were recorded at a sampling rate of 25 kHz. Contact impedance was <5 kΩ for all electrodes across all recording sessions.

EEG PREPROCESSING. The EEG data were preprocessed with Brain Vision Analyzer 2.0 (Brain Products) with the following steps. First, responses were off-line band-pass filtered from 80 to 1,000 Hz (12 dB/octave, zero phase shift). Responses were then segmented into epochs of 345 ms (−45 ms before stimulus onset and 300 ms after stimulus onset). After baseline correction of each trial to the mean voltage of the noise floor (−45 to 0 ms), trials with activity exceeding the range of ±35 μV were rejected. Across 3 days, for each recording session in *experiment 1*, 1,000 artifact-free FFR trials (500 for each polarity) were obtained for each tone, discarding any additional trials that might have been collected to match the number of trials across participants. In *experiment 2*, ~850 artifact-free FFR trials were obtained for each tone (425 for each polarity), across four recording sessions (pre-, mid-, and posttraining on *day 1*, and retention on *day 2*).

Pitch tracking analysis. Through the next series of analyses, we evaluated how accurately time-varying F0 information of the stimulus was tracked in the FFR with custom MATLAB scripts (The MathWorks, Natick, MA).

EXTRACTION OF F0 CONTOUR FROM FFR AND STIMULUS. First, the artifact-free FFR trials were averaged within polarities and then added across polarities to minimize stimulus-related artifact and cochlear microphonic contribution (Chandrasekaran and Kraus 2010). Next, we extracted the F0 contour from the averaged FFR waveform, using a sliding window (40 ms) autocorrelation-based procedure (see, e.g., Bidelman et al. 2011; Chandrasekaran et al. 2012; Krishnan et al. 2004; Krishnan et al. 2005; Skoe et al. 2014; Wong et al. 2007). The autocorrelation algorithm was developed by Boersma (1993). The procedure works as follows: A 40-ms sliding window was applied over the time course of the FFRs, which is assumed to encompass response from 10 to 260 ms. The 40-ms sliding window was shifted in 10-ms steps, to produce a total of 22 overlapping bins. To estimate the period of FFRs, the autocorrelation function was applied to each of the 22 time bins. The maximum (peak) autocorrelation value (ranging from −1 to 1) was searched over a lag value of 5–13.33 ms, a range that encompasses the time-variant periods of the F0 contours for *tone 1* and *tone 2*. The peak autocorrelation value as well as the corresponding lag were recorded for each bin. The lag was taken as the period of the F0 contour at each bin. To derive the F0 contour, the lag value associated with each of the 22 autocorrelation peaks was converted to frequency by taking it reciprocal. The resulting frequency values were concatenated to form a 22-point F0 contour. The same short-term autocorrelation algorithm was applied to the stimulus waveforms (from 0 to 250 ms) to derive the respective stimulus F0 contour (see Fig. 1*A* for F0 contours extracted from the stimuli, *tone 1* and *tone 2*).

EVALUATION OF PITCH TRACKING ACCURACY. For each EEG recording session from each participant, we calculated four metrics to assess the accuracy of the FFRs in tracking time-varying F0 information of the stimuli (*tone 1* and *tone 2*): stimulus-to-response correlation, peak autocorrelation, F0 error, and F0 amplitude. These metrics have been used as primary measures to evaluate the robustness of subcortical encoding of dynamic pitch patterns as a function of long-term language experience (see, e.g., Bidelman et al. 2011; Krishnan et al. 2005; Krishnan et al. 2009a; Krishnan et al. 2009b; Wong et al. 2007; Xu et al. 2006), as well as short-term auditory training (Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008). In this section, we describe how the four metrics of focus were calculated.

For the stimulus-to-response correlation metric (expressed as a value ranging from 0 to 1), we computed the normalized cross-correlation between F0 contours extracted from the stimuli and those derived from the corresponding FFRs (see, e.g., Bidelman et al. 2011; Krishnan et al. 2004; Krishnan et al. 2005). The maximum absolute cross-correlation coefficients were recorded as estimates of stimulus-to-response correlation. The purpose of using cross-correlation was to account for latency differences between F0 contours from the stimulus and the FFR. This metric represents the extent to which the F0 contours from the stimulus and the FFR are correlated, with a higher value indicating more faithful encoding of the stimulus' F0 in the FFR.

For the peak autocorrelation metric (expressed as a value between −1 to 1), we averaged the peak autocorrelation values across the 22 time bins in the FFR. Peak autocorrelation is a measure of the degree of periodicity in the FFR (Krishnan et al. 2010a), which indicates the robustness of neural phase-locking to the F0 of the stimulus (Carcagno and Plack 2011; see, e.g., Krishnan et al. 2004; Krishnan et al. 2005; Selinger et al. 2016; Wong et al. 2007). A larger value means a higher degree of periodicity in the FFR, suggesting more robust neural phase-locking to the stimulus F0.

For the F0 error metric, we first ran a point-by-point subtraction between F0 contours extracted from the stimuli and those derived from the corresponding FFRs across all 22 time bins. We took the absolute value of the differences, and averaged across the 22 time bins, to derive a single number that reflects the number of hertz that the F0 contour of FFR deviated, on average, from the stimulus F0 contour (Skoe et al. 2014; Song et al. 2008). A smaller value indicates higher faithfulness in the tracking of the stimulus F0 contour.

For the metric of F0 amplitude, we first applied fast Fourier transform (FFT) to each of the 22 time bins in the FFR. For each time bin, the raw amplitude of F0 was recorded as the FFT amplitude of the frequency that corresponds to the F0 detected by the above-mentioned short-term autocorrelation algorithm. The noise floor of F0 was also recorded as the average amplitude of a 50-Hz bin that is 5 Hz apart from the F0 (25-Hz bins on either side of the F0). For example, if the F0 in one time bin detected by the short-term autocorrelation algorithm is 102 Hz, the raw amplitude of F0 in that bin is defined as the amplitude of 102 Hz and the noise floor of F0 is defined as the combined average amplitude within 72–97 Hz and 107–132 Hz. For each of the 22 time bins, the raw amplitude of F0 was subtracted from the noise floor of F0. The average of such amplitude differences across the 22 bins was considered as the F0 amplitude. This metric quantifies the magnitude of spectral energy dedicated to the representation of F0 while accounting for the artificial boost in magnitude due to varying noise floors in each subject.

ADDITIONAL METRIC. We also computed another metric that does not directly relate to dynamic changes in pitch, i.e., signal-to-noise ratio (SNR) of the FFRs. To derive this metric, we divided the root mean square (RMS) amplitude of the FFR region (10–260 ms) of the FFR waveform by the RMS amplitude of the prestimulus region (−45 to 0 ms). This metric indicates the amount of neural activity in the FFR period relative to background noise (Russo et al. 2004).

Experiment 1: Long-Term Language Experience-Dependent Subcortical Plasticity to Linguistically Relevant Pitch Patterns

Participants. Adult native speakers of Mandarin Chinese (n = 10, 7 women; mean age = 26.3 yr, SD = 2.47 yr) and American English (n = 10, 6 women; mean age = 21.5 yr, SD = 3.24 yr) were recruited from the University of Texas at Austin. Native English participants reported no significant exposure to any tonal language. All participants had either no history of formal music training or no significant music expertise (<6 yr of continuous music training, not currently practicing). This inclusionary criterion was implemented because previous evidence has shown that music training influences the subcortical auditory processing of speech stimuli (Bidelman et al. 2011; Schön et al. 2004; Wong et al. 2007). All participants reported no previous history or current diagnosis of neuropsychiatric disorders, no use of neuropsychiatric medication, and no prior history of a hearing deficit. Hearing sensitivity for all participants was better than 20 dB hearing level (HL) for octave frequencies from 250 to 8,000 Hz.

EEG recording and analysis procedures. All participants completed repeated sessions of electrophysiological recording of FFRs to *tone 1* and *tone 2* across three consecutive days. The recording and analysis procedures are described in *General Materials and Methods*.

Statistical analysis. A three-way repeated-measures analysis of variance (ANOVA) was conducted on each of the four pitch-related metrics, as well as SNR. For metrics of stimulus-to-response correlation and peak autocorrelation, we first converted the correlational values into Fisher's *Z* scores to improve the normality of the data and then used the converted *Z* scores for statistical analyses (Wong et al. 2007). In each analysis, group (Chinese vs. English) was included as a between-subject factor and day (*1*, *2*, *3*) and stimulus (*tone 1* vs. *tone 2*) were included as within-subject factors. We report Greenhouse-Geisser-corrected results.

Experiment 2: Subcortical Plasticity to Linguistically Relevant Pitch Patterns After Short-Term Sound-to-Category Training

In this experiment, we examined the impact of a single session of sound-to-category training on the neural representation of linguistic pitch patterns in adolescents and young adults. To assess training-related subcortical plasticity, we restricted our analyses to the three metrics that demonstrated significant language experience-dependent plasticity in *experiment 1*, i.e., stimulus-to-response correlation, peak autocorrelation, and F0 error.

Participants. Twenty-two adolescents (10 female; mean age = 15.75 yr, SD = 1.12 yr) and eighteen adults (10 female; mean age = 20.73 yr, SD = 2.31 yr) took part in *experiment 2*. Two adolescent participants and three adult participants were excluded from analyses because of too few artifact-free trials (<850 trials for each tone on any session). All participants were monolingual native speakers of English and reported no significant exposure to tonal languages. Consistent with *experiment 1*, all participants reported no significant music expertise (<6 yr of continuous music training, not currently practicing), no previous history or current diagnosis of neuropsychiatric disorders, no use of neuropsychiatric medication, and no prior history of a hearing deficit. Hearing sensitivity for all participants was better than 20 dB HL for octave frequencies from 250 to 8,000 Hz.

Behavioral procedures. STIMULI. Stimuli for the sound-to-category training task were identical to those used in Chandrasekaran et al. (2014c). Four native Mandarin Chinese speakers (2 female) produced four Mandarin tones (*tone 1*, *tone 2*, *tone 3*, and *tone 4*) in citation form in the context of five monosyllabic words (*bu*, *di*, *lu*, *ma*, and *mi*). The 80 stimuli were normalized for RMS amplitude at 70 dB SPL and a duration of 440 ms. A novel set of 160 stimuli were generated for a retention-generalization task, using identical stimulus generation procedures in the sound-to-category training task. Specifically, eight new native Mandarin Chinese speakers (4 female) originally from China, produced the four Mandarin tones (*tone 1*, *tone 2*, *tone 3*, and *tone 4*) in citation form in the context of the same five monosyllabic words used in the training task. The 160 stimuli were normalized for RMS amplitude at 70 dB SPL and duration of 400 ms. Furthermore, all 240 stimuli (80 for the training task and 160 for the retention-generation task) were judged to be natural, with tone identification accuracy >95% by native speakers ($n = 5$).

SOUND-TO-CATEGORY TRAINING TASK. Participants were trained to categorize the four Mandarin tones over a session of sound-to-category training. As illustrated in Fig. 3*A*, each trial began with a fixation cross in the center of the screen for 750 ms. The stimulus was then presented binaurally via insert earphones (ER-3; Etymotic Research, Elk Grove Village, IL). Participants were instructed to categorize the stimulus into one of four categories by pressing the number keys (1, 2, 3, or 4) on a gamepad, corresponding to *tone 1*, *tone 2*, *tone 3*, and *tone 4*, respectively. No other instructions were provided. After their response, feedback was displayed for 1,000 ms. The response-to-feedback interval was fixed at 500 ms. The content of feedback varied depending on the accuracy of the response ("RIGHT" vs. "WRONG"). If the participant failed to generate a response within 2,000 ms after stimulus offset, a cautionary feedback ("TIME") was presented and the response did not register. These timed-out trials were coded as incorrect. All 80 stimuli were presented once within each of the six training blocks in a randomized sequence.

RETENTION-GENERALIZATION TASK. Participants completed a task that tested retention and generalization after a day following the training task. Similar to the sound-to-category training task, participants were instructed to categorize each stimulus into one of four categories by pressing the respective number keys on a keyboard (1, 2, 3, or 4). The stimuli utilized in this task were the 160 novel stimuli not encountered by participants during the sound-to-category training task. Each trial began with a fixation cross that appeared in the center of the screen for 750 ms, followed by one stimulus presented binaurally via Sennheiser HD 280 Pro circumaural headphones. Unlike the sound-to-category training task, the participants did not receive feedback at the end of each trial in this task. Participants had unlimited time to respond. Once they responded, the task moved on to the next trial. All 160 stimuli were presented once in a randomized sequence.

EEG recording and analysis procedures. On *day 1* (training day), we collected FFRs to Mandarin tones (*tone 1* and *tone 2*) before any training occurred (pretraining), after three blocks (80 trials/block) of sound-to-category training (midtraining), and immediately after another three blocks (80 trials/block) of sound-to-category training (posttraining). After ~24 h, participants returned to the laboratory to complete a final EEG recording (retention). This EEG recording was immediately followed by the behavioral retention-generalization sound-to-category task. The recording and analysis procedures are consistent with *General Materials and Methods*.

Statistical analysis. BEHAVIORAL TASKS. First, we examined learning (i.e., improvements across training blocks) in the sound-to-category training task. Here, we used mixed-effects modeling implemented in the lme4 package, using a binomial logit link (Bates et al. 2012) in R (R Development Core Team 2013). This type of analysis models mixed-effects logistic regression, where the estimates of the model output correspond to the log odd or probability of producing a correct response. Learning was assessed via three mixed-effects models for all four tones (overall accuracy) and also restricted to *tones 1* and *2*. These were the tones that were used in the EEG portion of the experiment. For these analyses, the responses of participants were coded on each trial as "correct" or "incorrect," with the timed-out trials also coded as "incorrect." This trial-by-trial accuracy was treated as the dependent variable, where the estimates corresponded to the log odds of producing a correct response. The fixed effects of interest were block (*1–6*; 80 trials/block), group (adult vs. adolescent, with adult as the reference level), and their interaction terms. By-subject random intercept was included as a random effect. We have employed similar analyses methods in previous publications involving sound-to-category training (Chandrasekaran et al. 2014c; Chandrasekaran et al. 2016; Reetzke et al. 2016; Yi et al. 2016).

Second, we examined the extent to which learning was retained and generalized to new talkers after a day of no training. A two-way repeated-measures ANOVA was conducted. Group (adult vs. adolescent) was included as a between-subject factor, and session [pretraining (*block 1*), posttraining (*block 6*) vs. generalization-retention] was included as a within-subject factor. We report Greenhouse-Geisser-corrected results. We applied this analysis to overall categorization performance (i.e., all 4 tones) as well as *tone 1* and *tone 2* categorization performance, respectively.

FFR METRICS. A three-way repeated-measures ANOVA was conducted to examine the effect of short-term sound-to-category training on the neural representation of linguistic pitch patterns. In this analysis, group (adolescent vs. adult) was included as a between-subject factor and FFR recording session (pre-, mid-, and posttraining and retention) and stimulus (*tone 1* vs. *tone 2*) were included as within-subject factors. We examined the three metrics that showed significant language experience-dependent plasticity in *experiment 1*, i.e., stimulus-to-response correlation, peak autocorrelation, and F0 error. Similar to *experiment 1*, we converted stimulus-to-response correlation and peak autocorrelation into Fisher's *Z* scores to improve the normality of the data and used the converted *Z* scores for statistical analyses (Wong et al. 2007). We report Greenhouse-Geisser corrected results.

RESULTS

Experiment 1

Figure 1*B* displays FFR waveforms to *tone 1* and *tone 2* recorded from an example Chinese participant and an example English participant across the 3 test days. The morphology of FFRs was highly consistent across days within participants.

We employed three-way (group, day, and stimulus) repeated-measures ANOVA to test the effects of language experience and test day on the four pitch tracking metrics: stimulus-to-response correlation, peak autocorrelation, F0 error, and F0 amplitude.

For the stimulus-to-response correlation metric, we found significant main effects of group [$F(1,18) = 5.607$, $P = 0.029$, $\eta_p^2 = 0.238$] and stimulus [$F(1,18) = 34.905$, $P < 0.001$, $\eta_p^2 = 0.66$] and a significant interaction effect between group and stimulus [$F(1,18) = 5.974$, $P = 0.025$, $\eta_p^2 = 0.249$]. The main effect of day, as well as the other two- and three-way interaction effects between the factors, did not reach statistical significance (all $P > 0.48$, $\eta_p^2 < 0.04$). Pooling across days, post hoc analyses revealed that stimulus-to-response correlation was significantly higher in the Chinese group relative to the English group for *tone 2* [Chinese vs. English: 0.725 (SD = 0.224) vs. 0.56 (SD = 0.16); $t(45.072) = 3.218$, $P = 0.002$, uncorrected] but not for *tone 1* [Chinese vs. English: 0.468 (SD = 0.11) vs. 0.452 (SD = 0.101); $t(58) = 0.635$, $P = 0.528$, uncorrected] (Fig. 2*A*).

For the peak autocorrelation metric, we found a significant main effect of group [$F(1,18) = 4.411$, $P = 0.05$, $\eta_p^2 = 0.197$], indicating that FFRs in the Chinese group exhibited a higher degree of periodicity relative to the English group [Chinese vs. English: 0.701 (SD = 0.097) vs. 0.634 (SD = 0.101)] (Fig. 2*B*). The main effect of day or stimulus was not significant (both $P > 0.21$, η_p^2 ranging from 0.066 to 0.188). None of the two-way or three-way interaction effects between group, day, and stimulus reached significance (all $P > 0.07$, η_p^2 ranging from 0.021 to 0.25).

For the F0 error metric, we found a significant main effect of group [$F(1,18) = 7.706$, $P = 0.012$, $\eta_p^2 = 0.300$], indicating that FFRs in the Chinese group had smaller errors in representing stimulus F0 contour compared with the English group [Chinese vs. English: 4.694 (SD = 3.162) vs. 7.548 (SD = 4.76)] (Fig. 2*C*). We also found a significant main effect of stimulus [$F(1,18) = 6.983$, $P = 0.017$, $\eta_p^2 = 0.28$], suggesting that FFRs to *tone 2* exhibited more errors in representing the F0 contour of the stimulus compared with FFRs to *tone 1* [*tone 1* vs. *tone 2*: 5.457 (SD = 4.215) vs. 6.785 (SD = 4.258)] (Fig. 2*C*). The main effect of day was not significant ($P = 0.284$, $\eta_p^2 = 0.068$). None of the two-way or three-way interaction effects between group, day, and stimulus reached significance (all $P > 0.19$, η_p^2 ranging from 0.033 to 0.088).

For the metrics of F0 amplitude and SNR, there were no significant main effects of group, day, or stimulus (all $P > 0.17$, η_p^2 ranging from 0.005 to 0.063). None of the two-way or three-way interactions between group, day, and stimulus reached significance (all $P > 0.088$, η_p^2 ranging from 0.002 to 0.134).

The main effect of group, paired with the absence of a significant main effect of day and related interaction effects across three of the pitch tracking metrics, suggests that long-term language experience is highly stable across multiple days of testing. To substantiate this interpretation of the findings we further examined the main effect of day for each of the five metrics, using a Bayesian model selection approach (Masson 2011). Specifically, we transformed the sums of squares from the ANOVAs to generate Bayesian information criterion probabilities (pBIC) associated with the occurrence of the null hypothesis ($H_{0,}$ i.e., stability in FFR metrics across test days) and the alternative hypothesis (H_1, i.e., change in FFR metrics across test days). According to the Raftery (1995) criterion, there was strong evidence in favor of the null hypothesis across the metric of stimulus-to-response correlation [$p(H_0|D) = 0.957$], peak autocorrelation [$p(H_0|D) = 0.964$], F0 amplitude [$p(H_0|D) = 0.963$], and SNR [$p(H_0|D) = 0.959$]. For the F0 error metric, we found weak evidence in favor of the null hypothesis [$p(H_0|D) = 0.543$]. For each of the five metrics, there was not even weak evidence in favor of the alternative hypothesis [all $p(H_1|D) < 0.5$].

From these results, we conclude that pitch tracking of linguistic pitch patterns is more accurate and robust for native Chinese listeners compared with native English listeners and that this language experience-dependent effect is not modulated by repeated passive exposure and in turn remains stable across multiple days.

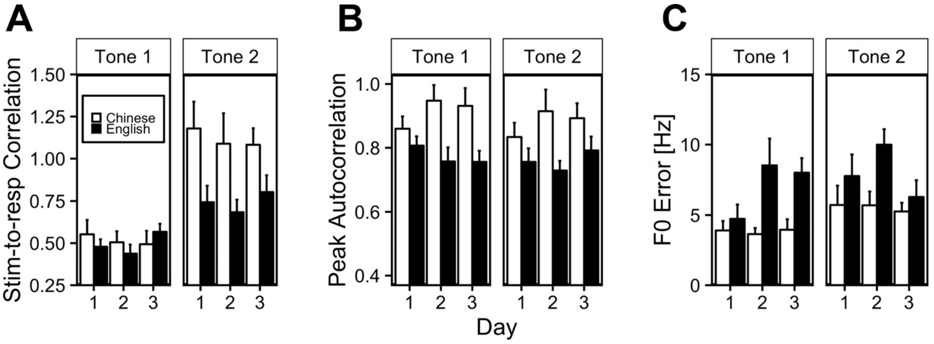

Fig. 2. Pitch-tracking accuracy for Chinese and English groups across 3 days of passive exposure in *experiment 1*. Results are plotted separately for *tone 1* and *tone 2*. *A*: mean stimulus-to-response correlation, Fisher's *Z*-transformed. *B*: mean peak autocorrelation, Fisher's *Z*-transformed. *C*: mean fundamental frequency error (F0 error). Error bars denote 1 SE.

Experiment 2

Behavioral tasks. LEARNING IN THE SOUND-TO-CATEGORY TRAINING TASK. First, we examined learning (i.e., improvements across training blocks) in the sound-to-category training task across all four tones. For the adolescent group, the mean accuracy improved from 32.1% (SD = 14.4%) in the first block to 61.2% (SD = 0.9%) in the final block. For the adult group, the mean accuracy improved from 31.4% (SD = 16.0%) in the first block to 59.2% (SD = 11.2%) in the final block.

The mixed-effects analysis revealed that the group × block interaction was not significant [$b = 0.01$, SE = 0.02, $Z = 0.80$, $P = 0.42$], suggesting that learning rate as a function of block was similar between adolescents and adults. The simple effect of block was significant [$b = 0.26$, SE = 0.02, $Z = 17.10$, $P < 0.0001$], suggesting that the overall tone categorization performance for both groups significantly improved across blocks. The simple effect of group was not significant [$b = -0.08$, SE = 0.30, $Z = -0.27$, $P = 0.79$], indicating that before training the overall tone categorization performance was not significantly different between the two groups. Similar overall tone categorization performance between groups across all training blocks is illustrated through Fig. 3*D* group confusion matrices.

Next, we examined categorization accuracies specific to the *tone 1* category. As shown in Fig. 3*B*, the mean accuracy for the adolescent group was 33.0% (SD = 18.4%) in the first block and increased to 67.0% (SD = 13.5%) in the final block. The mean accuracy for the adult group was 32.7% (SD = 22.3%) in the first block and increased to 59.7% (SD = 17.6%) in the final block. Results from the mixed-effects analysis revealed that the group × block interaction was significant [$b = 0.09$, SE = 0.04, $Z = 2.14$, $P = 0.03$], suggesting that the learning rate of *tone 1* as a function of block was greater for the adolescent group relative to the adult group. The simple effect of block was significant [$b = 0.25$, SE = 0.03, $Z = 7.94$, $P < 0.0001$], suggesting that *tone 1* categorization performance for both groups significantly improved across blocks. The simple effect of group was not significant [$b = -0.35$, SE = 0.37, $Z = -0.96$, $P = 0.34$], indicating that before training *tone 1* categorization performance was not significantly different across the two groups.

Finally, we examined categorization accuracies specific to the *tone 2* category. As shown in Fig. 3*B*, for the adolescent group the mean accuracy improved from 27.3% (SD = 19.7%) in the first block to 54.3% (SD = 17.3%) in the final block. For the adult group the mean accuracy improved from 34.7% (SD = 14.7%) in the first block to 49.0% (SD = 17.5%) in the

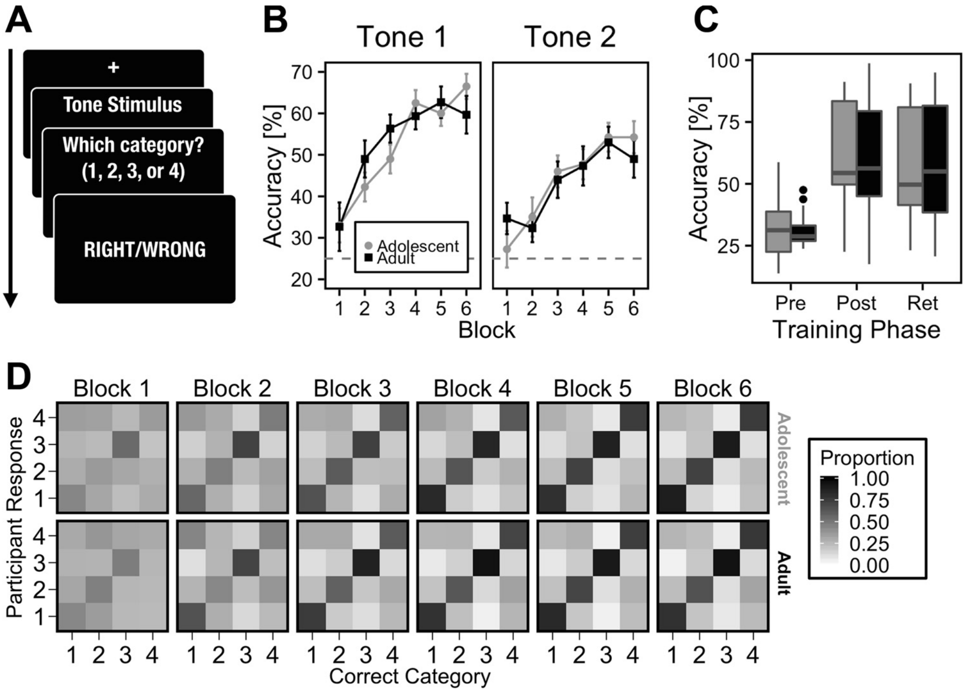

Fig. 3. Sound-to-category training paradigm and results in *experiment 2*. *A*: trial design for the sound-to-category training task. Each trial began with a fixation cross in the center of the screen for 750 ms. After that, a stimulus was presented for a fixed duration (440 ms). The participants were given a 2,500-ms time window to categorize the stimulus into 1 of 4 categories (*1*, *2*, *3*, or *4*). Corrective feedback (1,000 ms) was presented after participants' response. The response to feedback interval was fixed at 500 ms. *B*: overall group-by-block accuracy performance on the sound-to-category training task for the categorization of Mandarin *tone 1* and *tone 2*. *x*-Axis represents each learning block (*blocks 1–6*; 80 trials/block). *y*-Axis represents the accuracy of correct responses per block for adolescents and adults. Error bars denote 1 SE. *C*: retention and generalization results on the sound-to-category task: proportion of accurate responses for adolescents (gray) and adults (black) across training phases (pretrain: *block 1*; posttrain: *block 6*; and *day 2*: retention). Center line on each box plot denotes the median score, edges of the box denote the 25th and 75th percentiles, and whiskers extend to data points that lie within 1.5 times the interquartile range. Points outside this range denote outliers. *D*: tone category confusion matrices for adolescents (*top*) and adults (*bottom*) derived from the sound-to-category training response patterns. Each matrix corresponds to a training block; each column corresponds to participant category responses, and each row corresponds to the correct stimulus category. The shade of a given cell denotes the proportion of the category-response combination within a given block, averaged across all participants, ranging from high (black; value = 1.00) to low (white; value = 0.03). Higher values within the diagonal cells, extending from *bottom left* to *top right* corner of the matrix, correspond to correct responses; other cells denote errors. Informal visual inspection of the data suggests that overall sound-to-category training for both groups was associated with decreased confusion between *tone 1* and *tone 2*. This was confirmed by linear mixed-effects regression analyses demonstrating an improvement in categorization accuracy across blocks.

final block. Results from the mixed-effects analysis revealed that the group × block interaction was significant [$b = 0.09$, SE = 0.04, $Z = 2.20$, $P = 0.03$], suggesting that the learning rate of *tone 2* as a function of block was greater for the adolescent group relative to the adult group. The simple effect of block was significant [$b = 0.19$, SE = 0.03, $Z = 6.22$, $P < 0.0001$], suggesting that *tone 2* categorization performance for both groups significantly improved across blocks. The simple effect of group was not significant [$b = -0.30$, SE = 0.39, $Z = -0.78$, $P = 0.43$], indicating that before training *tone 2* categorization performance was not significantly different across the two groups.

RETENTION AND GENERALIZATION. We employed a two-way (group, session) repeated-measures ANOVA to examine overall retention-generalization performance relative to pre (*block 1*)- and post (*block 6*)-sound-to-category training sessions. As demonstrated in Fig. 3*C*, for overall accuracy, we found a significant main effect of session [F (2, 66) = 5.77, $P < 0.001$, $\eta_p^2 = 0.305$]. The main effect of group ($P = 0.99$, $\eta_p^2 = 0.00001$), as well as the two-way interaction effect between the group and session interaction ($P = 0.75$, $\eta_p^2 = 0.002$), did not reach statistical significance. Post hoc pairwise comparisons revealed that overall accuracy was significantly higher ($P < 0.001$, uncorrected) in post-sound-to-category training (*block 6*: mean = 60.32%; SD = 8.83%) and generalization-retention (mean = 57.32%; SD = 9.39%) sessions relative to the pre-sound-to-category training session (*block 1*: mean = 31.82%; SD = 13.86%). The comparison between the posttraining and generalization-retention sessions did not reach statistical significance ($P > 0.52$, uncorrected). Overall, these results demonstrate that, for both groups, learned tone categories were retained from *day 1* training and also generalized to novel talkers.

We conducted two additional two-way (group, session) repeated-measures ANOVA to examine *tone 1* and *tone 2* retention-generalization performance relative to pre (*block 1*)- and post (*block 6*)-sound-to-category training performance. For both *tone 1* and *tone 2* training accuracy, we found a significant main effect of session [*tone 1*: $F(2, 66) = 24.68$, $P \leq 0.001$, $\eta_p^2 = 0.223$; *tone 2*: $F(2, 66) = 2.58$, $P \leq 0.001$, $\eta_p^2 = 0.188$]. For both tones, the main effect of group [*tone 1*: $P = 0.84$, $\eta_p^2 = 0.0007$; *tone 2*: $P = 0.93$, $\eta_p^2 = 0.0002$], as well as the two-way interaction effect between the group and session interaction [*tone 1*: $P = 0.14$, $\eta_p^2 = 0.023$; *tone 2*: $P = 0.20$, $\eta_p^2 = 0.014$)] did not reach statistical significance. Post hoc pairwise comparisons revealed that overall accuracy was significantly higher [*tone 1*: $P < 0.001$, uncorrected; *tone 2*: $P <0 .001$, uncorrected] in the post-sound-to-category training [*tone 1 block 6*: mean = 63.57%, SD = 14.96%; *tone 2 block 6*: mean = 52.00%, SD = 14.68%] and generalization-retention [*tone 1*: mean = 55.50%, SD = 16.43%; *tone 2*: mean = 57.64%, SD = 10.69%] sessions relative to the pre-sound-to-category training session [*tone 1 block 1*: mean = 32.86%, SD = 18.63%; *tone 2 block 1*: mean = 30.43%, SD = 18.63%]. The comparison between the posttraining and generalization-retention sessions did not reach statistical significance [*tone 1*: $P = 0.17$, uncorrected; *tone 2*: $P = 0.32$, uncorrected]. Overall, these results demonstrate that for both groups the learned *tone 1* and *tone 2* categories were retained from *day 1* training and also generalized to novel talkers.

FFR metrics. We employed a three-way (group, session, and stimulus) repeated-measures ANOVA to examine the impact of training on neural representation of linguistic pitch patterns.

For the stimulus-to-response correlation metric, we found a significant main effect of stimulus [$F(1, 33) = 86.445$, $P < 0.001$, $\eta_p^2 = 0.724$], indicating more faithful encoding of the F0 of the stimulus in FFRs to *tone 2* relative to FFRs to *tone 1* [*tone 1* vs. *tone 2*: 0.416 (SD = 0.117) vs. 0.698 (SD = 0.19), $P = 0.029$]. We did not find a significant main effect of group ($P = 0.099$, $\eta_p^2 = 0.755$) or session ($P = 0.188$, $\eta_p^2 = 0.047$) or any interactions between the factors (all $P > 0.35$, η_p^2 ranging from 0.009 to 0.047).

For the peak autocorrelation metric, we found a significant main effect of session [$F(2.665, 87.931) = 2.952$, $P = 0.043$, $\eta_p^2 = 0.082$]. We found a significant session × group interaction [$F(2.665, 87.931) = 3.855$, $P = 0.015$, $\eta_p^2 = 0.105$]. We also found a significant main effect of stimulus [$F(1, 33) = 11.058$, $P = 0.002$, $\eta_p^2 = 0.251$], which suggests that FFRs to *tone 2* exhibited higher degree of periodicity relative to FFRs to *tone 1* [*tone 2* vs. *tone 1*: 0.64 (SD = 0.112) vs. 0.607 (SD = 0.119)]. We did not find a significant main effect of group ($P = 0.309$, $\eta_p^2 = 0.031$). We did not find any other significant interaction effects between the factors (all $P > 0.188$, η_p^2 ranging from 0.019 to 0.048).

Simple effect analysis demonstrated that for the adult group the effect of session was not significant [$F(2.478, 71.876) = 1.054$, $P = 0.365$, $\eta_p^2 = 0.035$]. For the adolescent group the effect of session was significant [$F(2.495, 97.308) = 11.916$, $P < 0.001$, $\eta_p^2 = 0.234$]. Pairwise comparisons revealed that peak autocorrelation significantly improved in the retention session (*day 2*) relative to the pretraining session (*day 1*) [retention vs. pretraining: 0.852 (SD = 0.191) vs. 0.783 (SD = 0.196), $P = 0.008$, uncorrected], to the midtraining session (*day 1*) [retention vs. midtraining: 0.852 (SD = 0.191) vs. 0.731 (SD = 0.18), $P < 0.001$, uncorrected], and to the posttraining session [retention vs. posttraining: 0.852 (SD = 0.191) vs. 0.748 (SD = 0.206), $P < 0.001$, uncorrected] (Fig. 4*B*). Furthermore, peak autocorrelation in the midtraining session significantly decreased relative to the pretraining session [midtraining vs. pretraining: 0.731 (SD = 0.18) vs. 0.783 (SD = 0.196), $P = 0.024$, uncorrected] (Fig. 4*B*). Other comparisons did not reach significance (both $P > 0.073$).

For the F0 error metric, we found a significant main effect of session [$F(2.618, 86.394) = 4.702$, $P = 0.006$, $\eta_p^2 = 0.125$]. We also found a main effect of stimulus [$F(1, 33) = 5.136$, $P = 0.03$, $\eta_p^2 = 0.135$], indicating more F0 errors for FFRs to *tone 1* relative to FFRs to *tone 2* [*tone 1* vs. *tone 2*: 7.609 (SD = 5.82) vs. 6.6 (SD = 4.138)]. There was no significant main effect of group ($P = 0.48$, $\eta_p^2 = 0.015$) or any interaction effects between the factors (all $P > 0.055$, η_p^2 ranging from 0.029 to 0.08). To understand the effect of session, pairwise comparison showed that F0 error was reduced in the retention session (*day 2*; see Fig. 4*C*) relative to the pretraining session (*day 1*) [retention vs. pretraining: 5.728 (SD = 4.126) vs. 6.919 (SD = 4.437), $P = 0.047$, uncorrected], to the midtraining session (*day 1*) [retention vs. midtraining: 5.728 (SD = 4.126) vs. 7.59 (SD = 4.939), $P < 0.001$, uncorrected], and to the posttraining session (*day 1*) [retention vs. posttraining: 5.728 (SD = 4.126) vs. 8.181 (SD = 6.267), $P < 0.001$, uncorrected]. Other comparisons did not reach significance (all $P > 0.059$).

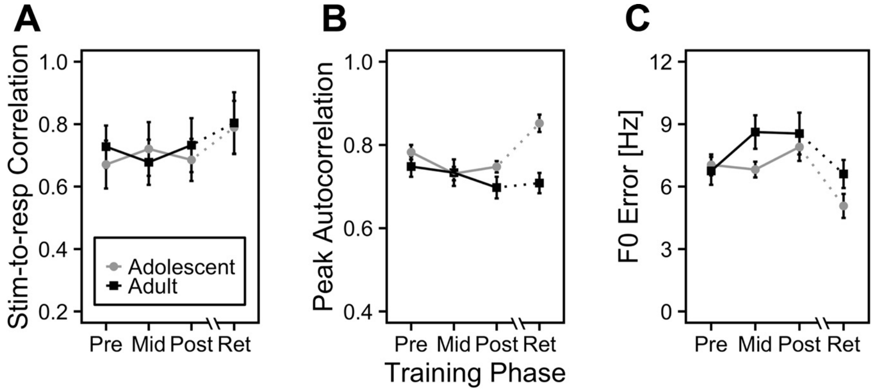

Fig. 4. Pitch tracking accuracy as a function of short-term sound-to-category training for adolescents and adults in *experiment 2*. Results reported here were collapsed across pitch tracking accuracy to *tones 1* and *2*. *A*: mean stimulus-to-response correlation, Fisher's *Z*-transformed. Panel illustrates negligible changes in stimulus-to-response correlation across the 4 sessions. *B*: mean peak autocorrelation, Fisher's *Z*-transformed. Panel illustrates an improvement in peak autocorrelation in the retention (Ret) session on *day 2* relative to the 3 sessions, pretraining (Pre), midtraining (Mid), and posttraining (Post), in *day 1* for adolescents but not for adults. *C*: mean fundamental frequency error (F0 error). Panel illustrates the decrease in F0 error in the retention (Ret) session relative to the 3 sessions on *day 1* for both groups. Error bars denote 1 SE.

Correlations between behavioral and FFR changes. Another way to examine effects of sound-to-category training on the subcortical pitch encoding is to assess the extent to which changes in sound-to-category training performance correlate with changes in FFR metrics. Pearson product-moment correlation coefficients were computed to examine the relationship between sound-to-category training gains (retention – pretraining (*block 1*)] and improvements in neural pitch tracking of linguistically relevant pitch patterns (retention – pretraining). Here we focused analyses only on the FFR metrics for which we found a significant main effect or interaction effect of session (i.e., peak autocorrelation and F0 error). Since there was a main effect of stimulus for both metrics, we examined the relationship between behavioral gains and changes in FFR metrics for each tone. First, we examined these correlations in the adolescent and adult groups separately. No correlation reached statistical significance for the adolescent group (all r ranged from -0.425 to 0.407, all $P > 0.062$) or for the adult group (r ranged from -0.330 to 0.203, all $P > 0.230$). Next, we conducted correlational analyses across all participants. When focusing on each stimulus individually, behavioral gains in *tone 2* significantly correlated with increase in peak autocorrelation to *tone 2* ($r = 0.368$, $P = 0.030$) as well as decrease in F0 error to *tone 2* ($r = -0.347$, $P = 0.041$). No other correlations reached statistical significance (all r ranged from -0.060 to 0.079, all $P > 0.65$). The pooled correlations suggest that increases in neural pitch tracking may contribute directly to improvements in the learning of novel speech categories. We speculate that group differences might emerge with a larger sample size or with a larger age differential between the adolescent and adult groups. Future work is needed to test these hypotheses.

DISCUSSION

In *experiment 1*, we assessed the stability of long-term language experience-dependent plasticity in the neural encoding of linguistic pitch patterns over multiple days. In *experiment 2*, we examined the extent to which induced behavioral relevance led to subcortical plasticity for the encoding of linguistically relevant pitch patterns as reflected by FFR recordings before, during, and immediately after a single session of sound-to-category training as well as a day after training ceased. Additionally, we assessed the extent to which adolescents and adults differed on learning novel speech categories and subcortical plasticity associated with learning gains.

The results from *experiment 1* showed that long-term language experience modulated the subcortical encoding of linguistic pitch patterns, a finding that replicated prior work (e.g., Jeng et al. 2011b; Krishnan et al. 2005; Krishnan et al. 2009a; Krishnan et al. 2009b; Krishnan et al. 2010b; Xu et al. 2006). Critically, we showed that the effect of long-term language experience on subcortical pitch encoding is highly stable across multiple days of recording. Passive exposure to thousands of trials of linguistic pitch patterns did not enhance subcortical pitch encoding in native English-speaking participants for whom the linguistically relevant pitch patterns were not behaviorally relevant. In contrast, findings from *experiment 2* demonstrated that, after behavioral relevance was established through a single session of feedback-based sound-to-category training, enhancement in subcortical pitch encoding was observed, with some developmental differences noted. Behaviorally, both adolescents and adults demonstrated similar abilities to learn, generalize, and retain all four Mandarin pitch categories. After 24 h posttraining, we found decreases in pitch-tracking errors (F0 error) for both adolescents and adults. However, an increase in neural phase-locking to the stimulus' F0 (peak autocorrelation) was only observed for adolescent participants.

Stability of Language Experience-Dependent Neural Plasticity to Linguistically Relevant Pitch Patterns

Compared with the native English-speaking group, FFRs elicited from the native Chinese-speaking group showed higher stimulus-to-response correlation, higher peak autocorrelation, and lower F0 error, especially for the more dynamic pitch pattern (i.e., *tone 2*). Both stimulus-to-response correlation and F0 error are metrics that are often used to assess the fidelity of FFR in the representation of stimulus F0 (stimulus-to-response correlation: e.g., Chandrasekaran et al. 2012; Krishnan et al. 2005; F0 error: e.g., Skoe et al. 2014; Song et al. 2008). Higher stimulus-to-response correlation and lower F0 error in native Chinese listeners, relative to native English listeners, demonstrates a more faithful subcortical tracking of F0 contours of the linguistic pitch patterns. The peak autocorrelation metric measures the degree of periodicity in the FFR (Krishnan et al. 2010a), which reflects the strength of neural phase-locking to the stimulus' F0 contour (Carcagno and Plack 2011; e.g., Krishnan et al. 2004; Krishnan et al. 2005; Selinger et al. 2016; Wong et al. 2007). Therefore, a higher value in this metric for Chinese listeners compared with native English listeners suggests more robust phase-locking to linguistic pitch patterns' F0 contours in Chinese listeners. To summarize, as evidenced by group differences on these primary pitch tracking metrics, there is a stronger representation of dynamic pitch patterns in Chinese listeners relative to native English listeners. These results are consistent with prior work that has used these same metrics in a similar cross-language approach (e.g., Jeng et al. 2011b; Krishnan et al. 2005; Krishnan et al. 2009a; Krishnan et al. 2009b; Krishnan et al. 2010b; Xu et al. 2006). A novel contribution of the present study is that we demonstrated the high stability of language experience-dependent neuroplasticity across multiple days of testing.

Superior neural encoding of dynamic pitch patterns in adult Chinese listeners (relative to English listeners) has been attributed to reorganization within the subcortical auditory circuitry as a result of extensive exposure to dynamic linguistic pitch patterns during language development (Jeng et al. 2011b; Krishnan et al. 2012; Krishnan and Gandour 2009). This local reorganization account is supported by animal models, which have shown that local reorganization of neural circuitry selectively enhances the encoding of stimulus parameters that are characteristic of one's auditory environment during development (Keuroghlian and Knudsen 2007; Knudsen 1999; Knudsen 2002). For example, in barn owls the development of midbrain representations of sound localization cues is guided by spatial information provided by vision. When the spatial information from vision is modified by prismatic spectacles, there is a local fine-tuning of auditory subcortical neurons to sound localization cues to match the altered visually spatial information (Brainard and Knudsen 1993). Relevant to the present study, this acquired reorganization at early stages of development persists into adulthood (Bajo et al. 2010; Linkenhoker et al. 2005). In other words, early experience-induced local changes are highly stable in an adult auditory system. In line with these findings, our results demonstrated high stability of early, long-term language experience-dependent neural plasticity to linguistic pitch patterns.

Neuroplasticity to Linguistically Relevant Pitch Patterns: Passive Exposure vs. Active Training

The role of passive exposure vs. active training in mediating neuroplasticity has been widely debated in the literature. In the visual domain, Watanabe and colleagues demonstrated that passive exposure leads to enhanced processing of the visual features in the human early visual cortices (i.e., V1; Watanabe et al. 2002a; Watanabe et al. 2002b). Given that the subcortical auditory regions are likened to the V1 in the nature of stimulus processing (Nelken 2004), we hypothesized that passive exposure to dynamic linguistic pitch patterns may change the neural encoding of these stimuli. Inconsistent with this prediction, in *experiment 1* we found no evidence of plastic changes in FFR metrics after passive exposure over multiple days.

In mature animal models, shifts in tuning properties of single neurons in the midbrain (i.e., evidence of plastic changes) are significantly greater when auditory stimulation is rendered behaviorally relevant relative to passive, repetitive sound exposure (see, e.g., Yan and Suga 1998). Consistent with these findings, studies that have compared plasticity between young and adult animals have suggested that the capacity for plasticity is quite restricted in adulthood (Brainard and Knudsen 1998). Unlike young animals, auditory repetition alone appears to induce negligible influence on stimulus representation in adult animals. Rather, to induce plasticity, the stimulus must become behaviorally relevant to the adult animal (for review see Keuroghlian and Knudsen 2007). In a series of studies, Suga and colleagues demonstrated that the corticofugal (top-down) modulatory system plays an essential role in the mediation of subcortical and cortical plasticity to behaviorally relevant signals in adults animals (for review see Suga 2008; Suga 2012; Suga et al. 2002). In line with these findings, in *experiment 2* we demonstrated plastic changes in FFR metrics that followed a single session of active sound-to-category training. Importantly, improvements in FFR metrics were not observed during or immediately after the training task (i.e., in the midtraining or posttraining session); rather, improvements in neural pitch tracking only emerged after participants were brought back on a second day after training (i.e., in the generalization-retention session). Moreover, the changes in FFR metrics after training were associated with behavioral gains in the sound-to-category training task. Training-related changes in *experiment 2* did not appear to be as robust as the stable long-term language experience-dependent plasticity observed in *experiment 1*. Taking these findings together, we speculate that the results of the present study largely refute the possibility that training-related plasticity is a simple result of exposure.

Why is the sound-to-category training paradigm effective in inducing neuroplasticity? There are at least two factors in the training paradigm that may explain the result of neuroplasticity. First, the learning of sound-to-category mapping is guided by reinforcements via feedback. Successful learning of sound-to-category mapping has been found to activate the reward-based circuitry, including the caudate, putamen, and the ventral striatum (Yi et al. 2016). Second, the present paradigm involved a multitalker training paradigm, which allowed the participant to selectively attend or tune to more reliable features across talkers (i.e., pitch direction) (Chandrasekaran et al. 2014c) that are more relevant for successful categorization of the linguistic pitch patterns (Chandrasekaran et al. 2016). Reinforcements and selective attention are critical elements in determining neuroplasticity in sensory processing in adulthood (Roelfsema et al. 2010). The neuroplastic changes in sensory processing are circumscribed to features most relevant to the task (Roelfsema et al. 2010). In accordance with this argument, we observed a correlation between behavioral and FFR changes only for *tone 2*, which has a more dynamic pitch pattern relative to *tone 1*.

The impact of active training over multiple days on neural representation of dynamic pitch patterns has been previously demonstrated (e.g., Chandrasekaran et al. 2012; Skoe et al. 2014; Song et al. 2008). However, it is unclear whether plastic changes in human adolescents and adults require extensive training over multiple days to years, or if change can occur after training over a shorter duration. Numerous animal studies suggest the potential of training-induced plasticity at a shorter timescale (e.g., 30 min; Gao and Suga 2000), at least at the level of single neurons (e.g., Gao and Suga 1998; Gao and Suga 2000; Yan and Suga 1998). Consistent with animal studies, *experiment 2* demonstrated that an hour-long single session of training results in auditory plasticity for both adolescents and adults. Beyond this, for the first time we showed that short-term training-induced plasticity, as indexed by FFR metrics, emerged only a day (~24 h) after training.

The establishment of training-induced plasticity can be interpreted in the context of the predictive tuning (PT) model (Chandrasekaran et al. 2014b). According to the PT model, in a mature system the auditory cortices continuously fine-tune the subcortical encoding of incoming auditory signals via corticofugal pathways. The corticofugal tuning operates by constantly anticipating the incoming stimulus. When the prediction is accurate, subcortical representation of the stimulus is enhanced (Chandrasekaran et al. 2014b). In the case of *experiment 2*, after sound-to-category training we posit that cortical representation of learned dynamic pitch patterns is strengthened, as suggested by a previous functional neuroimaging study using the sound-to-category training task (Yi et al. 2016). Improved cortical representation facilitates the prediction of incoming dynamic pitch patterns via corticofugal pathways. Therefore, after training, subcortical representation of dynamic pitch patterns becomes more robust. Intriguingly, we found that while there was an immediate improvement in behavioral gains for sound-to-category categorization performance, subcortical pitch encoding did not emerge until after a 24-h period.

The later emergence of training-induced subcortical plasticity may reflect a period of sleep-dependent consolidation after training (Alain et al. 2015). According to the reverse hierarchical theory (Ahissar et al. 2009), we speculate that the training-induced plasticity in subcortical encoding emerges later because there is a hierarchical order for sound-to-category training to influence sensory representation, with training effects likely appearing earlier on higher levels (e.g., auditory cortical areas) of neural presentation and later on lower levels (e.g., auditory subcortical regions). It should be noted that these gains in neural pitch tracking observed in the present study may not persist and relate to learning over longer periods of time (e.g., weeks to months).This argument is consistent with recent animal work suggesting that learning-related neural changes emerge within the initial days of learning and may assist learning (Reed et al. 2011), yet these neural changes do not persist and are not necessary for continued stable performance of learned behavior (Molina-Luna et al. 2008; Reed et al. 2011). Future studies on the time course of experience-dependent neural plasticity are needed to better elucidate the dynamics of short-term training.

Adolescence and Auditory Subcortical Plasticity After Feedback-Based Sound-to-Category Learning

Adolescence is a transitional period in human development characterized by ongoing and often heightened changes in neural structure and function. This developmental plasticity is found in numerous brain regions, including the auditory system (Mahajan and McArthur 2012; Paus et al. 1999). Structurally, maturation of corticofugal and cortical-cortical connectivity is ongoing during adolescence (Paus et al. 1999). Functionally, continued maturation of subcortical and cortical auditory processing has also been demonstrated throughout adolescence (Albrecht et al. 2000; Bishop et al. 2007; Fitzroy et al. 2015; Krizman et al. 2015; Mahajan and McArthur 2012; Ponton et al. 2000; Skoe et al. 2015a). To the best of our knowledge, this is the first study to track the dynamics of subcortical plasticity after feedback-based sound-to-category learning in adolescents relative to adults.

According to the classic interpretation of the critical period hypothesis, language learning that occurs later in life is significantly limited because of neural commitment to the first language (Krashen 1973; Lenneberg et al. 1967; Snow and Hoefnagel-Höhle 1978). In turn, second language learning is expected to be challenging for adolescents and adults in the acquisition of novel speech categories. Contrary to this, the present study demonstrated that both adolescents and adults are able to learn, generalize, and retain novel linguistically relevant speech patterns well above chance level. Both groups showed similar overall sound-to-category accuracy and learning rate for the four Mandarin tone pitch patterns. The only developmental difference found behaviorally was in the learning rate for the two pitch speech categories (*tone 1* and *tone 2*) used to elicit FFRs. Like our study, a recent study on reinforcement learning (Davidow et al. 2016; DiMenichi and Tricomi 2016) demonstrated faster learning rates in adolescents relative to adults. These group differences were associated with greater activity in striatal learning neural networks. In line with this study, it is possible that heightened sensitivity of reward processing mediated by striatal learning systems in adolescents relative to adults facilitated the better learning rate for *tones 1* and *2*. This interpretation is further a previous study demonstrating that successful learning of sound-to-category mapping involves the reward-based circuitry, which includes the striatum (Yi et al. 2016).

The developmental differences noted through these two metrics, i.e., similar group changes for pitch tracking (F0 error), relative to different changes in pitch strength (peak autocorrelation) may indicate that the peak autocorrelation metric is more sensitive in capturing developmental differences in the neural representation of pitch patterns (Jeng et al. 2011a) after auditory training. The observed difference in the strength of neural phase-locking to the stimulus F0 contour (indexed by peak autocorrelation) between adolescents and adults may reflect a greater potential for plasticity at the level of subcortical nuclei in adolescence, due to ongoing maturation of corticofugal circuitry (Paus et al. 1999), and related serotonergic system (Grossmann et al. 2011; Hall et al. 2011; Sodhi and Sanders-Bush 2004).

The strength of neural phase-locking has been thought to reflect the number of auditory fibers that phase-lock to the period of the F0 of the evoking stimuli (Plack et al. 2014). Inhibitory and excitatory circuits of auditory subcortical structures have been implicated in the mediation of phase-locking changes through the recruitment of auditory fibers phase-locking to the stimulus period and the inhibition of auditory fibers firing at different periods (Burger and Pollak 1998; Pollak 1999). As shown in animal models, the serotonergic system innervates auditory subcortical nuclei densely (Hurley and Pollak 1999; Klepper and Herbert 1991) and, in turn, through excitatory-inhibitory interactions modulates auditory subcortical responses to sound stimuli (Hurley and Sullivan 2012). A recent human study evaluated the relationship between serotonin transporter expression and the strength of neural phase-locking to the F0 of a speech stimulus, as reflected by the peak autocorrelation of the FFR. They found that individuals with low serotonin transporter expression, relative to those with medium to high expression, exhibited a stronger neural phase-locking to the stimulus F0 (Selinger et al. 2016). Based on this previous body of literature, we speculate that ongoing maturation of subcortical circuitry and function in adolescence (Fitzroy et al. 2015; Mahajan and McArthur 2012; Paus et al. 1999; Ponton et al. 2000; Skoe et al. 2015a) relates to lower serotonin transporter expression, and in turn fine-tunes neural responses to periodicity (e.g., F0) of speech stimuli during this period of development. We add to this body of literature by demonstrating that behavioral relevance of stimuli may be needed for developmental differences in subcortical auditory responses to be observed. Further studies are needed to examine training-related developmental differences in subcortical auditory processing and the interaction between age and serotonin genetic variation in modulating subcortical auditory processing.

Only one study to date has investigated experience-dependent plasticity in adolescence after auditory training (Tierney et al. 2015). This longitudinal study demonstrated that auditory training in the form of 3 yr of in-school music training led to stable subcortical responses to the speech syllable /da/ that were thought to confer advantages on a test of phonological awareness (Tierney et al. 2015). Here we extend this previous work by assessing the extent to which a short period of sound-to-category training induces changes in the subcortical auditory processing in adolescents relative to adults. We found that adolescents and adults showed a similar pattern of a reduction in pitch-tracking errors (F0 error) to the linguistic pitch patterns after the sound-to-category training. Interestingly, we only observed an increase in pitch strength or the robustness of neural phase-locking to the stimulus F0 contour (peak autocorrelation) for adolescent participants.

Taken together, these results indicate that subcortical auditory plasticity, as reflected by the FFR, may be heightened during certain stages of human auditory development, such as adolescence. However, the fact that improvements in subcortical pitch tracking were also observed in adults indicates that subcortical auditory plasticity is not limited to a critical period of auditory development. It is important to note that the present study does not refute the possibility for greater behavioral learning gains and subcortical experience-dependent auditory plasticity in early childhood relative to later stages of development. There was only a 5-yr average age difference between the adolescent and adult groups. It is possible that more exaggerated developmental differences in neural pitch tracking might have been observed with a younger adolescent group and an older adult group. This notion is supported by the more robust subcortical encoding of linguistically relevant pitch patterns for early, long-term tonal language speakers in *experiment 1* relative to the training-related gains observed in *experiment 2*.

FFR Reflects Auditory Experiences

The present study combined with a body of prior human work suggests that the FFR can reflect auditory experiences across multiple timescales, such as long-term (years) language experience (see, e.g., Bidelman et al. 2011; Jeng et al. 2011b; Krishnan et al. 2005; Krizman et al. 2012), short-term (days to years) training (see, e.g., Carcagno and Plack 2011; Chandrasekaran et al. 2012; de Boer and Thornton 2008; Skoe et al. 2014; Song et al. 2008; Tierney et al. 2015), as well as a single session (minutes) of statistical learning (e.g., Skoe et al. 2013; Skoe et al. 2015b). The present study expands upon these findings in three ways. First, we provide evidence that long-term experience-dependent plasticity, as reflected by the FFR, demonstrates replicability (i.e., stability) across repeated days of recordings. Second, to our knowledge, the present study is one of the first to measure FFRs at different stages of learning (i.e., before, during, and immediately after) in a single-session training paradigm between adolescent and adult participants. This design allowed us to track the emergence of auditory plasticity as a function of training. We observed that training-related plasticity occurred only after behavioral relevance was established and consolidated. Third, the present study is the first to measure FFR after short-term training ceased (cf. cessation of long-term training, Skoe and Kraus 2012). We demonstrated that training-induced plasticity in subcortical processing, as indexed by FFR, emerged a day after training ceased. Taken together, these experiments suggest that the FFR is reliable in capturing the impact of auditory experiences spanning from minutes to years on auditory function as well as the dynamics of short-term training-related plasticity.

It is presumed that the site of plasticity, as reflected by FFR, is localized to subcortical auditory circuitry, with the inferior colliculus as a primary generator (Bidelman 2015; Chandrasekaran and Kraus 2010; King et al. 2016; Smith et al. 1975; Sohmer et al. 1977). However, a recent investigation using magnetoencephalography (MEG) suggests that F0 information in the FFR, the primary focus of many FFR studies including the present one, may be a result of collective activities from the subcortical nuclei and the auditory cortex (Coffey et al. 2016). These findings concur with an emerging perspective that the FFR reflects integrated phase-locked activity from the entire auditory system (Kraus and White-Schwoch 2015). The EEG method implemented here does not allow us to tease apart the subcortical and cortical contributions to FFR. Future studies may utilize neuroimaging methods like MEG to separate the sources of FFR and examine the plasticity of different FFR sources as a function of a variety of auditory experiences. One hypothesis is that long-term experiences may be more reflected in the subcortical sources, while short-term training-related effects may be more reflected in the cortical sources.

Conclusions

We present evidence that early, long-term language experience leaves a stable and lasting influence on the neural encoding of linguistic pitch patterns. Additionally, we demonstrate an enhancement of subcortical pitch encoding following induced behavioral relevance of linguistically relevant pitch patterns through a single session of speech sound-to-category training. Behaviorally, both adolescents and adults demonstrated similar abilities to learn, generalize, and retain all novel speech categories. Twenty-four hours after training, we found better subcortical pitch tracking for all participants, as evidenced by decreases in pitch-tracking errors. However, an increase in robustness of neural phase-locking was only observed for the adolescent group. Taking these findings together, we argue that that behavioral relevance is a critical component for the observation of plasticity in the subcortical encoding of pitch.

ACKNOWLEDGMENTS

We thank Whitney Barlow, John Parsons, and the rest of the SoundBrain laboratory research assistants for help with participant recruitment and data collection. We thank Semeon Mesfin for assistance in generating the sound-to-category training task and Seth Koslov for generating the retention-generalization task. We thank Dr. Erika Skoe for providing the MATLAB codes to create the spectrograms and implement the pitch tracking analysis. We thank Han-Gyol Yi for providing comments on the manuscript.

GRANTS

This work was supported by the National Institute on Deafness and Other Communication Disorders (Grant R01 DC-013315 to B. Chandrasekaran) and an Undergraduate Research Fellowship to Whitney Barlow (mentored by B. Chandrasekaran and R. Reetzke). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Z.X. and R.R. performed experiments; Z.X., R.R., and B.C. analyzed data; Z.X., R.R., and B.C. interpreted results of experiments; Z.X., R.R., and B.C. prepared figures; Z.X., R.R., and B.C. drafted manuscript; Z.X., R.R., and B.C. edited and revised manuscript; Z.X., R.R., and B.C. approved final version of manuscript.

REFERENCES

Ahissar M, Nahum M, Nelken I, Hochstein S. Reverse hierarchies and sensory learning. *Philos Trans R Soc Lond B Biol Sci* 364: 285–299, 2009. doi:10.1098/rstb.2008.0253.

Alain C, Zhu KD, He Y, Ross B. Sleep-dependent neuroplastic changes during auditory perceptual learning. *Neurobiol Learn Mem* 118: 133–142, 2015. doi:10.1016/j.nlm.2014.12.001.

Albrecht R, Suchodoletz W, Uwer R. The development of auditory evoked dipole source activity from childhood to adulthood. *Clin Neurophysiol* 111: 2268–2276, 2000. doi:10.1016/S1388-2457(00)00464-8.

Asher JJ, García R. The optimal age to learn a foreign language. *Mod Lang J* 53: 334–341, 1969. doi:10.1111/j.1540-4781.1969.tb04603.x.

Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 13: 253–260, 2010. doi:10.1038/nn.2466.

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. *Cereb Cortex* 19: 2699–2707, 2009. doi:10.1093/cercor/bhp024.

Bates D, Maechler M, Bolker B. *lme4: Linear Mixed-Effects Models Using S4 Classes*. Vienna: R Project for Statistical Computing, 2012.

Bidelman GM. Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. *Hear Res* 323: 68–80, 2015. doi:10.1016/j.heares.2015.01.011.

Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 23: 425–434, 2011. doi:10.1162/jocn.2009.21362.

Bishop DV, Hardiman M, Uwer R, von Suchodoletz W. Maturation of the long-latency auditory ERP: step function changes at start and end of adolescence. *Dev Sci* 10: 565–575, 2007. doi:10.1111/j.1467-7687.2007.00619.x.

Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc Inst Phon Sci* 17: 97–110, 1993.

Brainard MS, Knudsen EI. Experience-dependent plasticity in the inferior colliculus: a site for visual calibration of the neural representation of auditory space in the barn owl. *J Neurosci* 13: 4589–4608, 1993.

Brainard MS, Knudsen EI. Sensitive periods for visual calibration of the auditory space map in the barn owl optic tectum. *J Neurosci* 18: 3929–3942, 1998.

Burger RM, Pollak GD. Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the inferior colliculus. *J Neurophysiol* 80: 1686–1701, 1998.

Carcagno S, Plack CJ. Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol* 12: 89–100, 2011. doi:10.1007/s10162-010-0236-1.

Chandrasekaran B, Koslov SR, Maddox WT. Toward a dual-learning systems model of speech category learning. *Front Psychol* 5: 825, 2014a. doi:10.3389/fpsyg.2014.00825.

Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47: 236–246, 2010. doi:10.1111/j.1469-8986.2009.00928.x.

Chandrasekaran B, Kraus N, Wong PC. Human inferior colliculus activity relates to individual differences in spoken language learning. *J Neurophysiol* 107: 1325–1336, 2012. doi:10.1152/jn.00923.2011.

Chandrasekaran B, Skoe E, Kraus N. An integrative model of subcortical auditory plasticity. *Brain Topogr* 27: 539–552, 2014b. doi:10.1007/s10548-013-0323-9.

Chandrasekaran B, Yi HG, Blanco NJ, McGeary JE, Maddox WT. Enhanced procedural learning of speech sound categories in a genetic variant of FOXP2. *J Neurosci* 35: 7808–7812, 2015. doi:10.1523/JNEUROSCI.4706-14.2015.

Chandrasekaran B, Yi HG, Maddox WT. Dual-learning systems during speech category learning. *Psychon Bull Rev* 21: 488–495, 2014c. doi:10.3758/s13423-013-0501-5.

Chandrasekaran B, Yi HG, Smayda KE, Maddox WT. Effect of explicit dimensional instruction on speech category learning. *Atten Percept Psychophys* 78: 566–582, 2016. doi:10.3758/s13414-015-0999-x.

Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7: 11070, 2016. doi:10.1038/ncomms11070.

Dahmen JC, Keating P, Nodal FR, Schulz AL, King AJ. Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron* 66: 937–948, 2010. doi:10.1016/j.neuron.2010.05.018.

Davidow JY, Foerde K, Galván A, Shohamy D. An upside to reward sensitivity: the hippocampus supports enhanced reinforcement learning in adolescence. *Neuron* 92: 93–99, 2016. doi:10.1016/j.neuron.2016.08.031.

de Boer J, Thornton AR. Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J Neurosci* 28: 4929–4937, 2008. doi:10.1523/JNEUROSCI.0902-08.2008.

Dean I, Harper NS, McAlpine D. Neural population coding of sound level adapts to stimulus statistics. *Nat Neurosci* 8: 1684–1689, 2005. doi:10.1038/nn1541.

DiMenichi BC, Tricomi E. Are you smarter than a teenager? Maybe not when it comes to reinforcement learning. *Neuron* 92: 1–3, 2016. doi:10.1016/j.neuron.2016.09.043.

Fitzroy AB, Krizman J, Tierney A, Agouridou M, Kraus N. Longitudinal maturation of auditory cortical function during adolescence. *Front Hum Neurosci* 9: 530, 2015. doi:10.3389/fnhum.2015.00530.

Gandour JT. Tone perception in far eastern languages. *J Phon* 11: 149–175, 1983.

Gandour JT, Harshman RA. Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Lang Speech* 21: 1–33, 1978. doi:10.1177/002383097802100101.

Gao E, Suga N. Experience-dependent corticofugal adjustment of midbrain frequency map in bat auditory system. *Proc Natl Acad Sci U S A* 95: 12663–12670, 1998. doi:10.1073/pnas.95.21.12663.

Gao E, Suga N. Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proc Natl Acad Sci U S A* 97: 8081–8086, 2000. doi:10.1073/pnas.97.14.8081.

Grossmann T, Johnson MH, Vaish A, Hughes DA, Quinque D, Stoneking M, Friederici AD. Genetic and neural dissociation of individual responses to emotional expressions in human infants. *Dev Cogn Neurosci* 1: 57–66, 2011. doi:10.1016/j.dcn.2010.07.001.

Hakuta K, Bialystok E, Wiley E. Critical evidence: a test of the critical-period hypothesis for second-language acquisition. *Psychol Sci* 14: 31–38, 2003. doi:10.1111/1467-9280.01415.

Hall IC, Sell GL, Hurley LM. Social regulation of serotonin in the auditory midbrain. *Behav Neurosci* 125: 501–511, 2011. doi:10.1037/a0024426.

Hornickel J, Knowles E, Kraus N. Test-retest consistency of speech-evoked auditory brainstem responses in typically-developing children. *Hear Res* 284: 52–58, 2012. doi:10.1016/j.heares.2011.12.005.

Hurley LM, Pollak GD. Serotonin differentially modulates responses to tones and frequency-modulated sweeps in the inferior colliculus. *J Neurosci* 19: 8071–8082, 1999.

Hurley LM, Sullivan MR. From behavioral context to receptors: serotonergic modulatory pathways in the IC. *Front Neural Circuits* 6: 58, 2012. doi:10.3389/fncir.2012.00058.

Jeng FC, Hu J, Dickman B, Lin CY, Lin CD, Wang CY, Chung HK, Li X. Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. *Int J Audiol* 50: 14–26, 2011a. doi:10.3109/14992027.2010.515620.

Jeng FC, Hu J, Dickman B, Montgomery-Reagan K, Tong M, Wu G, Lin CD. Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear Hear* 32: 699–707, 2011b. doi:10.1097/AUD.0b013e31821cc0df.

Johnson JS, Newport EL. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn Psychol* 21: 60–99, 1989. doi:10.1016/0010-0285(89)90003-0.

Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. *J Neurosci* 28: 4000–4007, 2008. doi:10.1523/JNEUROSCI.0012-08.2008.

Johnson MH, De Haan M. *Developmental Cognitive Neuroscience: An Introduction*. Hoboken, NJ: Wiley-Blackwell, 2015.

Keuroghlian AS, Knudsen EI. Adaptive auditory plasticity in developing and adult animals. *Prog Neurobiol* 82: 109–121, 2007. doi:10.1016/j.pneurobio.2007.03.005.

King A, Hopkins K, Plack CJ. Differential group delay of the frequency following response measured vertically and horizontally. *J Assoc Res Otolaryngol* 17: 133–143, 2016. doi:10.1007/s10162-016-0556-x.

Klepper A, Herbert H. Distribution and origin of noradrenergic and serotonergic fibers in the cochlear nucleus and inferior colliculus of the rat. *Brain Res* 557: 190–201, 1991. doi:10.1016/0006-8993(91)90134-H.

Knudsen EI. Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl. *J Comp Physiol A* 185: 305–321, 1999. doi:10.1007/s003590050391.

Knudsen EI. Instructed learning in the auditory localization pathway of the barn owl. *Nature* 417: 322–328, 2002. doi:10.1038/417322a.

Kral A, Eggermont JJ. What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res Brain Res Rev* 56: 259–269, 2007. doi:10.1016/j.brainresrev.2007.07.021.

Kral A, Hartmann R, Tillein J, Heid S, Klinke R. Congenital auditory deprivation reduces synaptic activity within the auditory cortex in a layer-specific manner. *Cereb Cortex* 10: 714–726, 2000. doi:10.1093/cercor/10.7.714.

Kral A, Hartmann R, Tillein J, Heid S, Klinke R. Hearing after congenital deafness: central auditory plasticity and sensory deprivation. *Cereb Cortex* 12: 797–807, 2002. doi:10.1093/cercor/12.8.797.

Kral A, Tillein J, Heid S, Hartmann R, Klinke R. Postnatal cortical development in congenital auditory deprivation. *Cereb Cortex* 15: 552–562, 2005. doi:10.1093/cercor/bhh156.

Krashen SD. Lateralization, language learning, and the critical period: some new evidence. *Lang Learn* 23: 63–74, 1973. doi:10.1111/j.1467-1770.1973.tb00097.x.

Kraus N, Skoe E, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann N Y Acad Sci* 1169: 543–557, 2009. doi:10.1111/j.1749-6632.2009.04549.x.

Kraus N, White-Schwoch T. Unraveling the biology of auditory learning: a cognitive-sensorimotor-reward framework. *Trends Cogn Sci* 19: 642–654, 2015. doi:10.1016/j.tics.2015.08.017.

Krishnan A, Bidelman GM, Gandour JT. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hear Res* 268: 60–66, 2010a. doi:10.1016/j.heares.2010.04.016.

Krishnan A, Gandour JT. The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain Lang* 110: 135–148, 2009. doi:10.1016/j.bandl.2009.03.005.

Krishnan A, Gandour JT, Bidelman GM. Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport* 23: 498–502, 2012. doi:10.1097/WNR.0b013e328353764d.

Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport* 20: 408–413, 2009a. doi:10.1097/WNR.0b013e3283263000.

Krishnan A, Gandour JT, Smalt CJ, Bidelman GM. Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain Lang* 114: 193–198, 2010b. doi:10.1016/j.bandl.2010.05.004.

Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J Cogn Neurosci* 21: 1092–1105, 2009b. doi:10.1162/jocn.2009.21077.

Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 25: 161–168, 2005. doi:10.1016/j.cogbrainres.2005.05.004.

Krishnan A, Xu Y, Gandour JT, Cariani PA. Human frequency-following response: representation of pitch contours in Chinese tones. *Hear Res* 189: 1–12, 2004. doi:10.1016/S0378-5955(03)00402-7.

Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc Natl Acad Sci U S A* 109: 7877–7881, 2012. doi:10.1073/pnas.1201575109.

Krizman J, Tierney A, Fitzroy AB, Skoe E, Amar J, Kraus N. Continued maturation of auditory brainstem function during adolescence: A longitudinal approach. *Clin Neurophysiol* 126: 2348–2355, 2015. doi:10.1016/j.clinph.2015.01.026.

Lenneberg EH, Chomsky N, Marx O. *Biological Foundations of Language*. New York: Wiley, 1967.

Linkenhoker BA, von der Ohe CG, Knudsen EI. Anatomical traces of juvenile learning in the auditory system of adult barn owls. *Nat Neurosci* 8: 93–98, 2005. doi:10.1038/nn1367.

Luna B, Padmanabhan A, O'Hearn K. What has fMRI told us about the development of cognitive control through adolescence? *Brain Cogn* 72: 101–113, 2010. doi:10.1016/j.bandc.2009.08.005.

Maddox WT, Chandrasekaran B. Tests of a dual-systems model of speech category learning. *Biling (Camb Engl)* 17: 709–728, 2014. doi:10.1017/S1366728913000783.

Maddox WT, Chandrasekaran B, Smayda K, Yi HG. Dual systems of speech category learning across the lifespan. *Psychol Aging* 28: 1042–1056, 2013. doi:10.1037/a0034969.

Mahajan Y, McArthur G. Maturation of auditory event-related potentials across adolescence. *Hear Res* 294: 82–94, 2012. doi:10.1016/j.heares.2012.10.005.

Malmierca MS, Cristaudo S, Pérez-González D, Covey E. Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29: 5483–5493, 2009. doi:10.1523/JNEUROSCI.4153-08.2009.

Masson ME. A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behav Res Methods* 43: 679–690, 2011. doi:10.3758/s13428-010-0049-5.

Molina-Luna K, Hertler B, Buitrago MM, Luft AR. Motor learning transiently changes cortical somatotopy. *Neuroimage* 40: 1748–1754, 2008. doi:10.1016/j.neuroimage.2007.11.018.

Nelken I. Processing of complex stimuli and natural scenes in the auditory cortex. *Curr Opin Neurobiol* 14: 474–480, 2004. doi:10.1016/j.conb.2004.06.005.

Paus T, Zijdenbos A, Worsley K, Collins DL, Blumenthal J, Giedd JN, Rapoport JL, Evans AC. Structural maturation of neural pathways in children and adolescents: in vivo study. *Science* 283: 1908–1911, 1999. doi:10.1126/science.283.5409.1908.

Plack CJ, Barker D, Hall DA. Pitch coding and pitch processing in the human brain. *Hear Res* 307: 53–64, 2014. doi:10.1016/j.heares.2013.07.020.

Pollak GD. Analysis of the role of inhibition in shaping responses to sinusoidally amplitude-modulated signals in the auditory system. *J Acoust Soc Am* 105: 932, 1999. doi:10.1121/1.426302.

Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clin Neurophysiol* 111: 220–236, 2000. doi:10.1016/S1388-2457(99)00236-9.

Pujol J, Deus J, Losilla JM, Capdevila A. Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology* 52: 1038–1043, 1999. doi:10.1212/WNL.52.5.1038.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2013. [document freely available at http://www/r-project org, 2015.]

Raftery AE. Bayesian model selection in social research. *Sociol Methodol* 25: 111–163, 1995. doi:10.2307/271063.

Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, Kilgard MP. Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron* 70: 121–131, 2011. doi:10.1016/j.neuron.2011.02.038.

Reetzke R, Maddox WT, Chandrasekaran B. The role of age and executive function in auditory category learning. *J Exp Child Psychol* 142: 48–65, 2016. doi:10.1016/j.jecp.2015.09.018.

Robinson BL, Harper NS, McAlpine D. Meta-adaptation in the auditory midbrain under cortical influence. *Nat Commun* 7: 13442, 2016. doi:10.1038/ncomms13442.

Roelfsema PR, van Ooyen A, Watanabe T. Perceptual learning rules based on reinforcers and attention. *Trends Cogn Sci* 14: 64–71, 2010. doi:10.1016/j.tics.2009.11.005.

Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. *Clin Neurophysiol* 115: 2021–2030, 2004. doi:10.1016/j.clinph.2004.04.003.

Schneider W, Eschman A, Zuccolotto A. *E-Prime: User's Guide*. Pittsburgh, PA: Psychology Software, 2002.

Schön D, Magne C, Besson M. The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41: 341–349, 2004. doi:10.1111/1469-8986.00172.x.

Selinger L, Zarnowiec K, Via M, Clemente IC, Escera C. Involvement of the serotonin transporter gene in accurate subcortical speech encoding. *J Neurosci* 36: 10782–10790, 2016. doi:10.1523/JNEUROSCI.1595-16.2016.

Skoe E, Chandrasekaran B, Spitzer ER, Wong PC, Kraus N. Human brainstem plasticity: the interaction of stimulus probability and auditory learning. *Neurobiol Learn Mem* 109: 82–93, 2014. doi:10.1016/j.nlm.2013.11.011.

Skoe E, Kraus N. A little goes a long way: how the adult brain is shaped by musical training in childhood. *J Neurosci* 32: 11507–11510, 2012. doi:10.1523/JNEUROSCI.1949-12.2012.

Skoe E, Krizman J, Anderson S, Kraus N. Stability and plasticity of auditory brainstem function across the lifespan. *Cereb Cortex* 25: 1415–1426, 2015a. doi:10.1093/cercor/bht311.

Skoe E, Krizman J, Spitzer E, Kraus N. The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience* 243: 104–114, 2013. doi:10.1016/j.neuroscience.2013.03.009.

Skoe E, Krizman J, Spitzer E, Kraus N. Prior experience biases subcortical sensitivity to sound patterns. *J Cogn Neurosci* 27: 124–140, 2015b. doi:10.1162/jocn_a_00691.

Smayda KE, Chandrasekaran B, Maddox WT. Enhanced cognitive and perceptual processing: a computational basis for the musician advantage in speech learning. *Front Psychol* 6: 682, 2015. doi:10.3389/fpsyg.2015.00682.

Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol* 39: 465–472, 1975. doi:10.1016/0013-4694(75)90047-4.

Snow CE, Hoefnagel-Höhle M. The critical period for language acquisition: evidence from second language learning. *Child Dev* 49: 1114–1128, 1978. doi:10.2307/1128751.

Sodhi MS, Sanders-Bush E. Serotonin and brain development. *Int Rev Neurobiol* 59: 111–174, 2004. doi:10.1016/S0074-7742(04)59006-2.

Sohmer H, Pratt H, Kinarti R. Sources of frequency following responses (FFR) in man. *Electroencephalogr Clin Neurophysiol* 42: 656–664, 1977. doi:10.1016/0013-4694(77)90282-6.

Song JH, Nicol T, Kraus N. Test-retest reliability of the speech-evoked auditory brainstem response. *Clin Neurophysiol* 122: 346–355, 2011. doi:10.1016/j.clinph.2010.07.009.

Song JH, Skoe E, Wong PC, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20: 1892–1902, 2008. doi:10.1162/jocn.2008.20131.

Steinberg L. A social neuroscience perspective on adolescent risk-taking. *Dev Rev* 28: 78–106, 2008. doi:10.1016/j.dr.2007.08.002.

Suga N. Role of corticofugal feedback in hearing. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 194: 169–183, 2008. doi:10.1007/s00359-007-0274-2.

Suga N. Tuning shifts of the auditory system by corticocortical and corticofugal projections and conditioning. *Neurosci Biobehav Rev* 36: 969–988, 2012. doi:10.1016/j.neubiorev.2011.11.006.

Suga N, Xiao Z, Ma X, Ji W. Plasticity and corticofugal modulation for hearing in adult animals. *Neuron* 36: 9–18, 2002. doi:10.1016/S0896-6273(02)00933-9.

Tierney AT, Krizman J, Kraus N. Music training alters the course of adolescent auditory development. *Proc Natl Acad Sci U S A* 112: 10062–10067, 2015. doi:10.1073/pnas.1505114112.

Watanabe T, Náñez JE Sr, Koyama S, Mukai I, Liederman J, Sasaki Y. Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nat Neurosci* 5: 1003–1009, 2002a. doi:10.1038/nn915.

Watanabe T, Sasaki Y, Náñez JE, Koyama S, Mukai I, Hibino H, Tootell RB. Psychophysics and fMRI reveal V1 as the locus of passive learning. *J Vis* 2: 557, 2002b. doi:10.1167/2.7.557.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10: 420–422, 2007.

Xu Y, Krishnan A, Gandour JT. Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport* 17: 1601–1605, 2006. doi:10.1097/01.wnr.0000236865.31705.3a.

Yan J, Suga N. Corticofugal modulation of time-domain processing of biosonar information in bats. *Science* 273: 1100–1103, 1996. doi:10.1126/science.273.5278.1100.

Yan W, Suga N. Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat Neurosci* 1: 54–58, 1998. doi:10.1038/255.

Yi HG, Maddox WT, Mumford JA, Chandrasekaran B. The role of corticostriatal systems in speech category learning. *Cereb Cortex* 26: 1409–1420, 2016. doi:10.1093/cercor/bhu236.

Zhang Y, Suga N, Yan J. Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387: 900–903, 1997. doi:10.1038/43180.