LANGUAGE EXPERIENCE SHAPES PROCESSING OF PITCH RELEVANT INFORMATION IN THE HUMAN BRAINSTEM AND AUDITORY CORTEX: ELECTROPHYSIOLOGICAL EVIDENCE

Ananthanarayan Krishnan, PhD[1]; Jackson T. Gandour, PhD[1]

[1]Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907, USA

rkrish@purdue.edu

Pitch is a robust perceptual attribute that plays an important role in speech, language, and music. As such, it provides an analytic window to evaluate how neural activity relevant to pitch undergo transformation from early sensory to later cognitive stages of processing in a well coordinated hierarchical network that is subject to experience-dependent plasticity. We review recent evidence of language experience-dependent effects in pitch processing based on comparisons of native vs. nonnative speakers of a tonal language from electrophysiological recordings in the auditory brainstem and auditory cortex. We present evidence that shows enhanced representation of linguistically-relevant pitch dimensions or features at both the brainstem and cortical levels with a stimulus-dependent preferential activation of the right hemisphere in native speakers of a tone language. We argue that neural representation of pitch-relevant information in the brainstem and early sensory level processing in the auditory cortex is shaped by the perceptual salience of domain-specific features. While both stages of processing are shaped by language experience, neural representations are transformed and fundamentally different at each biological level of abstraction. The representation of pitch relevant information in the brainstem is more fine-grained spectrotemporally as it reflects sustained neural phase-locking to pitch relevant periodicities contained in the stimulus. In contrast, the cortical pitch relevant neural activity reflects primarily a series of transient temporal neural events synchronized to certain temporal attributes of the pitch contour. We argue that experience-dependent enhancement of pitch representation for Chinese listeners most likely reflects an interaction between higher-level cognitive processes and early sensory-level processing to improve representations of behaviorally-relevant features that contribute optimally to perception. It is our view that long-term experience shapes this adaptive process wherein the top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural responses to specific behaviorally-relevant attributes of the stimulus. A theoretical framework for a neural network is proposed involving coordination between local, feedforward, and feedback components that can account for experience-dependent enhancement of pitch representations at multiple levels of the auditory pathway. The ability to record brainstem and cortical pitch relevant responses concurrently may provide a new window to evaluate the online interplay between feedback, feedforward, and local intrinsic components in the hierarchical processing of pitch relevant information.

INTRODUCTION

Pitch is an essential perceptual attribute in the processing of language and music [1, 2]. Functional brain-imaging studies provide strong evidence for hierarchical processing of pitch [3], starting in subcortical structures [4] (Griffiths, Uppenkamp, Johnsrude, Josephs & Patterson (2001) and continuing up through Heschl's Gyrus on to the planum polare and planum temporale [5-7]. Therefore, pitch provides an excellent window for studying experience-dependent effects on both cortical and brainstem components of a well-coordinated, hierarchical processing network. It is our view that for a complete understanding of the neural organization of language it is necessary to treat these processes as a set of hierarchical computations that are applied to representations at different stages (subcortical and cortical) along the processing hierarchy. Such representations, particularly of linguistically relevant features or dimensions, in turn are shaped by experience within a specific domain.

Indeed, recent empirical data show that these neural representations of pitch, at both the brainstem and cortical level, are shaped by one's experience with language and music [8-19] While it is not known how language/music experience shapes subcortical and cortical stages of pitch processing, it is likely that the neural processes underlying such experience-dependent plasticity at each stage along the processing hierarchy are modulated by a coordinated interplay between ascending, descending and local neural pathways which involve both sensory and cognitive components [20]. That is, feedback from language–dependent cortical processes shape early sensory level processing at both the brainstem and cortical level. These enhanced sensory level outputs transform later functionally more salient cortical representations that drive processes mediating linguistic performance. This review is largely confined to crosslanguage (Mandarin vs English) electrophysiological studies evaluating processing of linguistically-relevant pitch contours in the brainstem and auditory cortex. Based on empirical evidence, we propose a theoretical framework that includes local, feedback, and feedforward components to account for experience-dependent enhanced representations of pitch at brainstem and cortical stages of processing.

WHY THE FOCUS ON TONE LANGUAGES?

Tone languages are especially useful for studying pitch because variations in pitch convey part of the meaning of a word [21,22]. It is well established that dynamic variations in voice fundamental frequency (F0) provide the dominant acoustic cue for tonal recognition [23-25]. Mandarin Chinese is a lexical tone language with four phonologically distinctive tones (see Figure 1) characterized by syllable-level fundamental frequency (f0) pitch contours. These pitch contours are commonly described as high-level (T1, e.g., *yi1* "clothing"), high-rising (T2, e.g., *yi2* "aunt"), falling-rising (T3, e.g., *yi3* "chair"), and high-falling (T4, *yi4* "easy"). Such languages are to be distinguished from those in which pitch variations are usually not contrastive at the syllable or word level (e.g., English). In nontone languages however, variations in pitch may be used to signal stress and intonation patterns at post-lexical levels of representation. Thus, tone languages not only give us a physiologic window to evaluate how neural representations of linguistically-relevant pitch attributes emerge along the early stages of sensory processing in the hierarchy, but they may also shed light on the nature of interaction between early sensory levels and later higher levels of cognitive processing in the human brain.

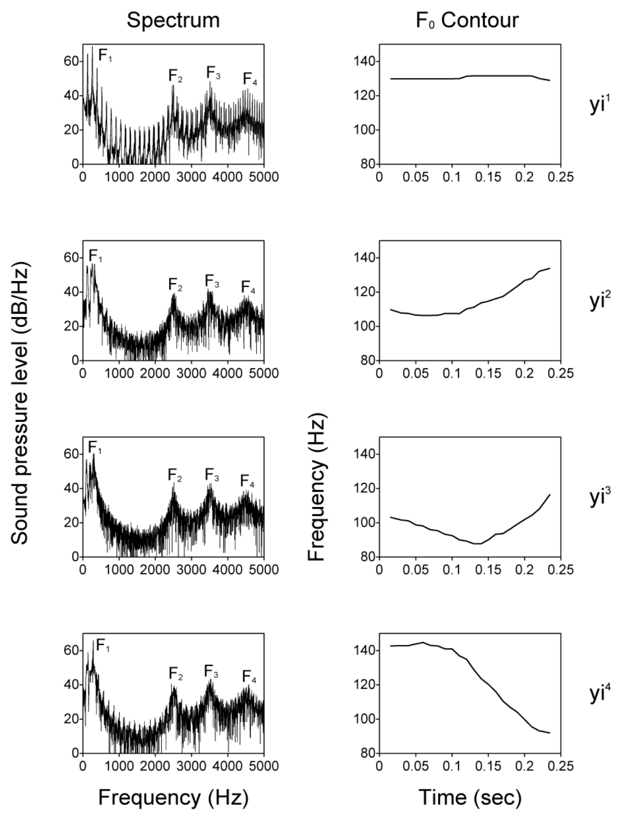


Figure 1. Spectra (left panel) and pitch contours of the four Mandarin lexical tones. Note that that were speech stimuli with invariant spectra across the four tones. The syllable is identified to the right of each panel.

LANGUAGE-EXPERIENCE SHAPES REPRESENTATION OF PITCH RELEVANT INFORMATION IN THE BRAINSTEM

Electrophysiological recordings are not only crucial for investigating questions about the hierarchy of pitch processing in the cerebral cortex, but also in subcortical structures [26] We present here empirical evidence suggesting that long-term language experience enhances the neural representation of linguistically relevant pitch in the human brainstem, well before evoked neural activity relevant to pitch is detected in the auditory cortex. Indeed, neural representation of pitch relevant attributes, as reflected in the scalp-recorded frequency following response (FFR), may emerge as early as 6-9 ms after stimulus onset [27]. In contrast, the pitch related neural activity in the auditory cortex emerges at about 140-170 ms post stimulus onset [28-32].

Characteristics of the brainstem frequency following response (FFR)

We use the FFR as a physiologic window into the early stages of subcortical processing relevant to pitch. The FFR reflects sustained phase-locked activity in a population of neural elements within the rostral brainstem, presumably the inferior colliculus [20, 33]. These responses can be easily recorded between scalp electrodes placed at high forehead and the seventh cervical vertebra (C7). The response is characterized by a waveform that follows the periodicities contained in both the envelope and temporal fine structure of complex sounds (see Figure 2).

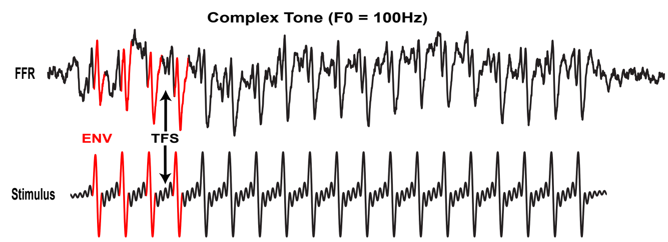


Figure 2. Frequency following responses (top trace) elicited by a complex sound (bottom trace). The response is characterized by neural phase-locking to the slower envelope periodicity (shown in red for a few cycles) and the faster temporal fine structure (TFS) of the complex sound (indicated by black arrow). Stimulus onset is shifted to the right to achieve temporal match between the stimulus and the response.

The temporal and spectral characteristics of complex sounds preserved in the FFR, can be extracted by frequency domain spectral and time domain (see figure 3) autocorrelation analysis (a measure of, correlation between the original signal and temporally delayed versions of the signal yielding high correlation for periodicities harmonically related to the fundamental frequency), respectively [34,35]. FFRs preserve spectrotemporal information relevant to the spectra (see Figure 4) [36] and pitch of steady-state [37] and dynamic speech [34,35] and nonspeech stimuli (see Figure 6) [38-41]. Importantly, the pitch-relevant information preserved in the FFR is strongly correlated with perceptual measures of pitch salience [38, 42]. Pitch salience is a measure of the strength of the perceived pitch. These findings suggest acoustic features relevant to pitch are preserved in the temporal pattern of phase-locked neural activity in the brainstem. Gockel, Carlyon, Mehta & Plack [43] observed that FFRs recorded using frequency shifted complex tones presented monaurally did preserve pitch relevant information, but that this information at the brainstem level was similar to that measured in an auditory nerve model. Also, they failed to observe any pitch relevant information in the FFRs to dichotically presented three tone harmonic stimuli. These findings led them to conclude that there was no additional pitch relevant processing at the level of the brainstem. Several arguments may be presented to counter this inference. First, if the temporal code for pitch available at the brainstem level also utilizes autocorrelation to determine the global distribution of interspike intervals from the temporal pattern of neural activity across a population of neurons, it would necessarily share certain fundamental attributes of the same temporal code operating at the level of the auditory nerve. Second, it is not clear that the dichotic stimuli used in their experiments would produce the same pitch as when all harmonics are presented to the same ear. Even if it does, its salience would be quite weak. It is possible that the FFR related neural activity is not sufficiently robust to preserve the less salient pitch for this stimulus. In our own experience, we have failed to measure FFR correlates of the less salient dichotic Huggins pitch. Finally, these inferences cannot convincingly account for the experience-dependent effects reflected in the FFR that are sensitive to specific attributes of dynamic pitch contours.

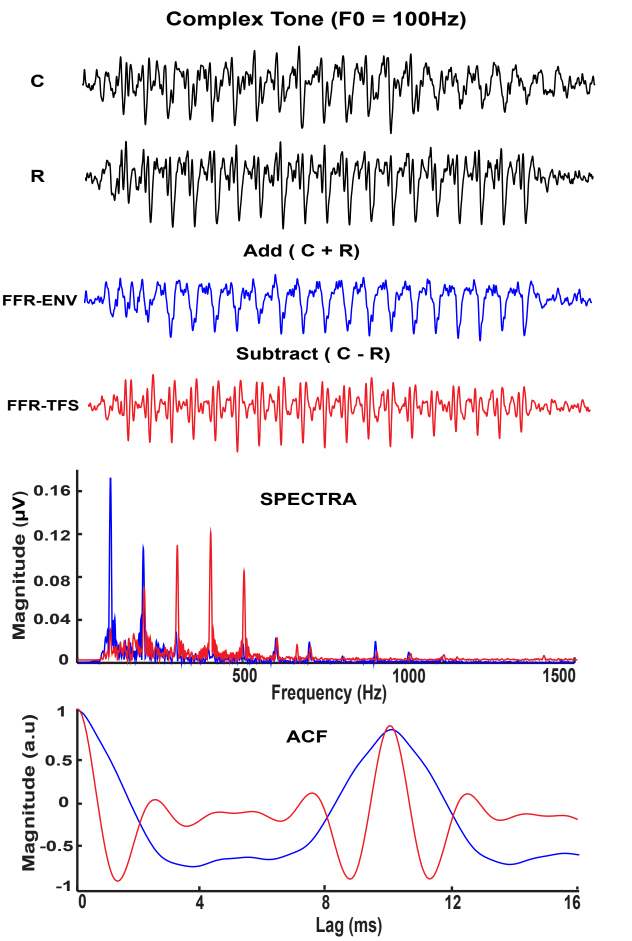


Figure 3. Frequency and time domain analysis of the FFR. The two top FFR waveforms (black) are responses to condensation (C) and rarefaction (R) onset stimulus polarity. The result of addition of these waveforms results in the FFR-ENV waveform (Blue), which is characterized by a prominent phase-locking to the envelope periodicity of the complex tone. In contrast, the subtracted (C-R) FFR shows phase locking to spectral components of the complex stimulus (Red). Frequency domain analysis (middle panel) shows that the envelope phase-locking, as expected, has a larger peak at F0 (blue) whereas phase locking to the spectral components of the stimulus are relatively smaller. The temporal analysis, (bottom panel) using autocorrelation, shows that for both responses there is a major peak at the fundamental periodicity of the complex tone.

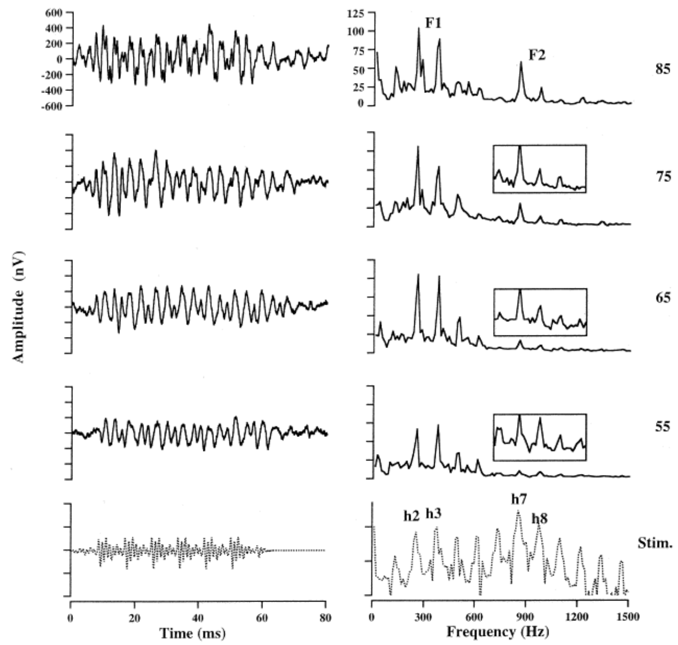


Figure 4. Grand averaged FFR waveforms and spectra are plotted as a function of stimulus level (dB nHL) for the English **vowel /u/**. The stimulus waveform and its spectrum (with F_1 and F_2 harmonics identified) are at the bottom of each panel. The amplified inset in the FFR spectral data clearly shows the F_2 harmonic peaks. Note the different amplitude scale for the stimulus spectrum. Note that the stimulus spectrum is without an amplitude scale and is used here to look for correspondence between the stimulus and the response spectrum. **Reprinted from Krishnan, 2003 with permission of Elsevier B.V.** A. Pitch tracking accuracy as a function of iteration steps. Note that pitch tracking accuracy increases more rapidly with increasing pitch salience for the Chinese listeners (red trace). B. Relationship between neural pitch strength and behavioral measure of pitch discrimination (F0 DL). Note that as F0 DL decreases with increasing pitch salience neural strength increases.

Evidence for experience-dependent plasticity in the human brainstem

The temporal pattern of the phase-locked neural activity generating the FFR preserves pitch-relevant information of lexical tones in both native (Mandarin) and nonnative (English) listeners. However, both pitch tracking accuracy (the strength of the correlation between the stimulus pitch contour, and the pitch contour extracted from the FFR) and periodicity strength (the strength of neural phase-locking to the fundamental periodicity of the stimulus, as reflected by the magnitude of the autocorrelation peak at a delay corresponding to the pitch period) is more robust in the former (Figure 5) because of long-term exposure to their native pitch contours [35].

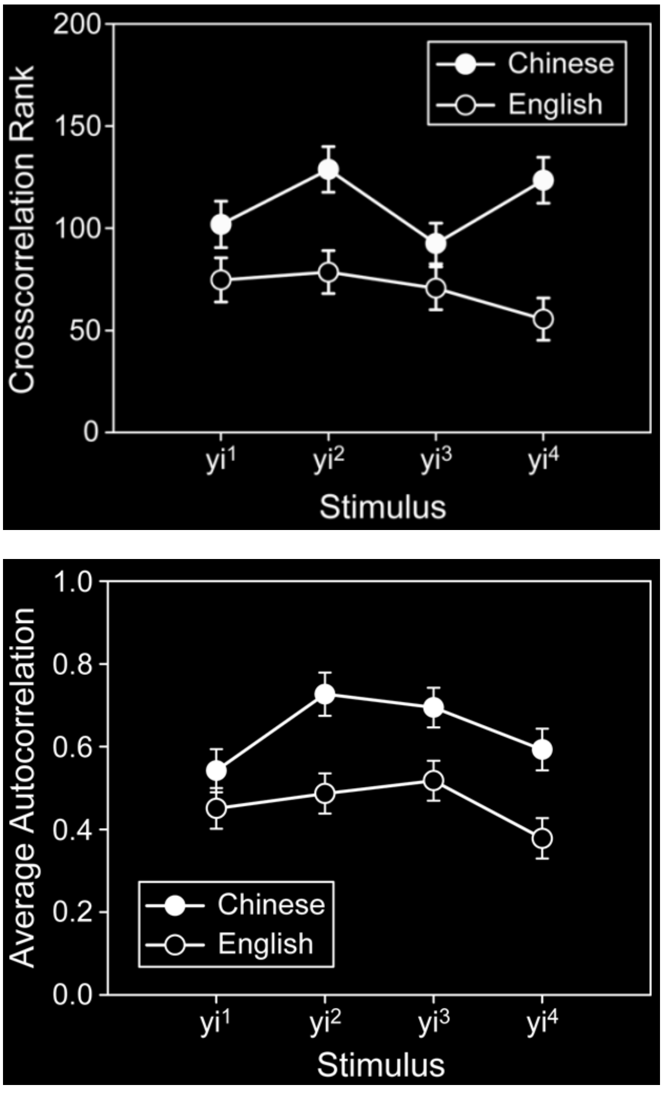
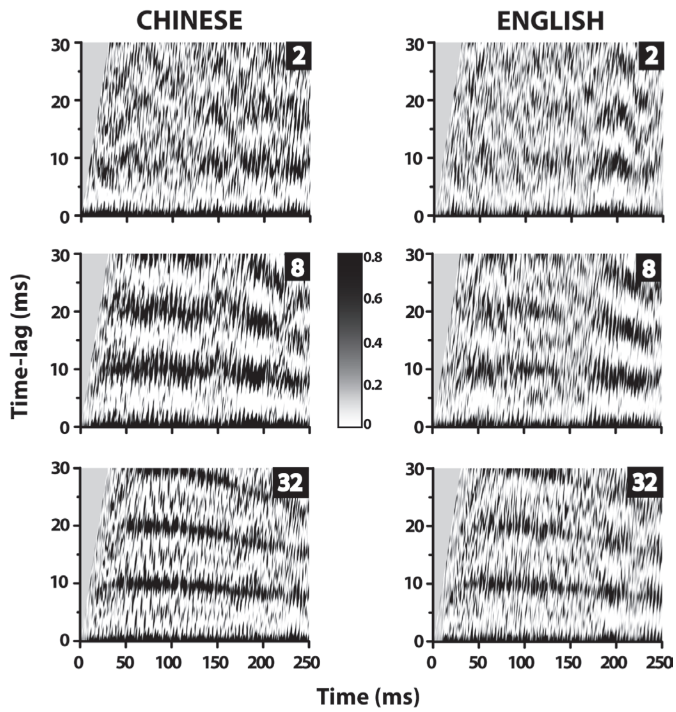


Figure 5. A & B. Comparison of pitch tracking accuracy (top panel) and pitch strength (bottom panel) for Chinese and English listeners for the four lexical tones. For both measure Chinese show better tracking accuracy and greater periodicity strength across the four stimuli compared to English listeners. Reprinted from Krishnan et al., 2005.

We therefore *conclude that early sensory representation of pitch-relevant information in the brainstem is sensitive to language experience*. This enhanced representation in Mandarin speakers is observed for native pitch contours whether they are presented in a speech or nonspeech context [35, 41]. Moreover, we find that Chinese exhibit more robust pitch representation precisely in those segments of Mandarin tones that contain rapid changes in pitch [40]. With respect to pitch features, brainstem representation of rising contours is found to be most important in discriminating Chinese from listeners of non-tone languages [44] Pitch features important to tone perception show increased resistance to degradation of the temporal regularity (decrease in temporal regularity reduces the perceived salience of pitch) of stimuli in Chinese listeners [45]. Iterated rippled noise (IRN) stimulus approximating the native lexical pitch contour of Tone 2 was used as a non-speech stimulus in this study.



Figures 6. Autocorrelograms derived from grand averaged FFR waveforms of Chinese (left) and English (right) groups in response to T2IRN at low (**n**=2), intermediate (**n**=8), and high (**n**=32) iteration steps. Little difference is seen between the Chinese and English group at low iteration steps (**n**=2; top row). By **n**=8 iterations, Chinese show clearer, tighter bands of temporal regularity (black) in FFR phase-locked activity at the fundamental period (1/f0) and its multiples as compared to the English group (middle row). At n=32, this superiority is even more pronounced (bottom row). Periodicity strength is indicated by the color gradient; darker shades indicate higher temporal regularity (i.e., more pronounced phase-locked activity). **Reprinted from Krishnan et al., 2010 with permission of Elsevier B.V.**

This IRN stimulus with a dynamic pitch contour was generated by applying a 32-step time-varying, delay-and-add iterative process to a broadband noise. These advantages in neural representation in the Mandarin listeners, however, diminish as the pitch acceleration within a pitch contour exceeds values beyond what occurs in natural speech [46]. Collectively, *these results lead us to conclude that experience-dependent neuroplasticity is primarily restricted to stimuli that are of functional relevance to the listener. Indeed, a strong correlation between neural and behavioral measures corroborates the view that neural representation of pitch-relevant information at a subcortical, sensory level of processing plays an important role in shaping tone perception [38].*

Language experience-dependent pitch processes in the brainstem are especially sensitive to the curvilinear shape of pitch contours that occur in natural speech. Enhanced representations do not occur for linear approximations of native rising or falling pitch contours [39, 47]. A nonnative, curvilinear pitch pattern similarly fails to elicit a language-dependent effect [39]. *Thus, experience-dependent effects in the brainstem are highly sensitive to specific fine-grained features of pitch patterns in one's native language.*

Finally, our comparisons between the language and music domains reveal overall enhancement in brainstem FFRs elicited by either musical or linguistic pitch patterns in musicians and tone language speakers alike [48,49]. Thus, long-term pitch experience seems to improve the brain's ability to represent pitch-relevant information regardless of the domain of expertise. However, subtle differences in these sensory representations suggest a domain-specific sensitivity to acoustic features that are part of the experience in each domain. Musicians, for example, show enhanced responses when pitch patterns intersect discrete notes along the musical scale; tone language speakers, on the other hand, during rapidly changing portions of tonal contours [48, 50]. Such cue weighting is consistent with the relative importance of these perceptual dimensions in their respective domains. *These findings collectively lead us to infer that both language and musical experience provide some mutual benefit to the neural representation of pitch-relevant information, but that specific features of the acoustic signal are highlighted in subcortical responses depending on their perceptual salience and function within a listener's domain of expertise.*

LANGUAGE-EXPERIENCE SHAPES REPRESENTATION OF PITCH RELEVANT INFORMATION IN THE AUDITORY CORTEX

Characteristics of the Pitch Onset Response (POR) using Magnetoencephalography (MEG)

At the cortical level, MEG has been used previously to study the sensitivity to pitch relevant periodicity, by investigating the N100m component. However, this component to a large extent simply represents sound onset and does not exclusively represent pitch [51-55]. To obtain pitch-specific response uncontaminated by the onset response, a novel stimulus paradigm was developed wherein two segments – an initial noise segment devoid of pitch evoked the onset responses only, followed temporally by an iterated rippled noise (IRN) segment, matched in intensity and overall spectral profile of the noise precursor, which elicited the pitch response components [28]. Interestingly, a transient pitch onset response (POR), with a latency of about 140-170 ms, was evoked from this noise-to-pitch transition. The reverse stimulus transition from pitch to noise failed to produce a POR. It has been proposed that the human POR, as measured by MEG, reflects synchronized cortical neural activity specific to pitch (28, 29; 56, 57]. POR latency and magnitude, for example, has been shown to depend on pitch salience. A more robust POR with shorter latency is observed for stimuli with stronger pitch salience as compared to those with weaker pitch salience. Source analyses [28, 52, 58] corroborated by human depth electrode recordings [31, 59] indicate that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing [5, 6, 60-62]. Given both its sensitivity to pitch and its salience, and consistency across a number of studies, the POR offers an excellent window for studying early sensory level representation of pitch specific information at the cortical level. Our preliminary POR data, extracted from scalp-recorded EEG, yielded *multiple* peaks (labeled as Pa (latency: 70-80 ms), Na (130-150 ms), Pb (200-215 ms), Nb (265-280 ms), and Pc (305-320 ms), where P and N represent peaks with positive and negative polarity, respectively) in addition to pitch onset response (POR). We therefore have chosen to designate this scalp-recorded neural activity as cortical pitch response (CPR).

Characteristics of the cortical pitch response (CPR) using electroencephalography (EEG)

Using a similar two-segment stimulus paradigm (see Figure 7), we demonstrated that the EEG-derived human cortical pitch response (CPR) elicited by IRN steady-state pitch stimuli increased in magnitude with increasing temporal regularity of the stimulus [32]. This change in CPR response amplitude with increasing stimulus temporal regularity was strongly correlated with behavioral measures of change in pitch salience (perceived strength of pitch). That the CPR is pitch specific is confirmed by its absence in response to a "no-pitch" IRN stimulus. Recently we examined the sensitivity of the multiple transient components of the CPR to specific temporal attributes associated with three (see Figures 8, 9 and 10), within-category variants of Mandarin Tone 2 (T2_150; T2_200; and T2_250) that varied in pitch acceleration (rate of change of fundamental frequency) and duration [63]. Our results showed that the latency of the pitch onset response (Na) was unaltered by changes in pitch acceleration (Figures 8 and 9). In contrast, response components Pa-Na, Na–Pb and Pb–Nb showed a systematic decrease in the interpeak latency, and a decrease in amplitude with increase in pitch acceleration (see figure 10 for amplitude changes with acceleration). These changes in interpeak latencies closely followed the time course of pitch change across the three stimuli. This is readily apparent if the stimulus pitch contours (in Figure 7) are compared to the 500-700 ms response window (in Figure 8). A strong correlation with pitch acceleration was observed for Na-Pb and Pb-Nb only - a putative index of pitch-relevant neural activity associated with the more rapidly-changing portions of the pitch contour. Pc–Nc marks unambiguously the stimulus offset. These data lead us to propose that in the early stages of cortical sensory processing of pitch, a series of neural markers flag different temporal attributes of a dynamic pitch contour: onset of temporal regularity (Na); changes in temporal regularity between onset and offset (Na–Pb, Pb–Nb); and offset

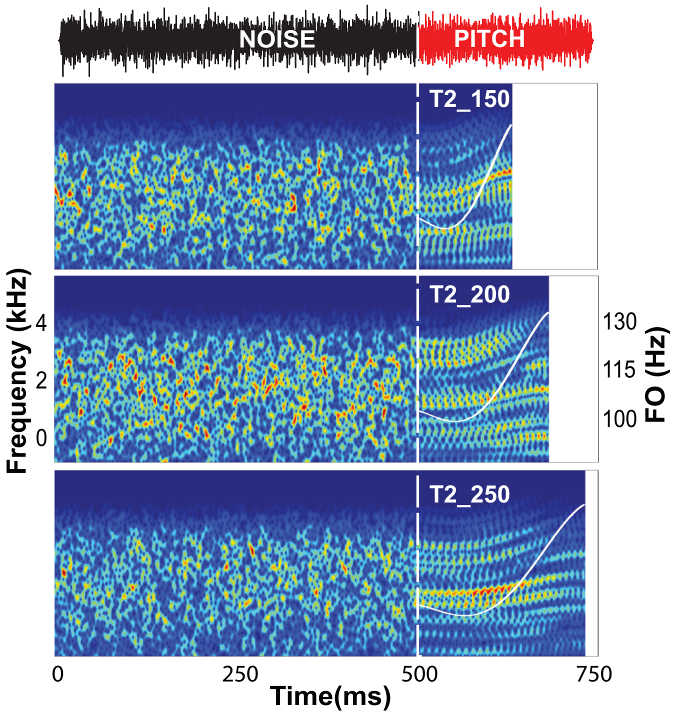


Figure 7. Noise-to-pitch Waveform (T2_250) and spectrograms of each of the three iterated rippled noise (IRN) stimulus conditions (T2_150; T2_200; T2_250) illustrate the experimental paradigm used to acquire brainstem and cortical responses concurrently. The vertical dashed line at 500 ms demarcates the transition from the initial noise segment to the final pitch segment. FFRs and CPRs were extracted from evoked responses beginning with the onset of the pitch. F0 contours (white) are superimposed on their respective pitch segment spectrograms. Within the pitch segment, the waveform (top) shows robust periodicity at a high IRN iteration step (n=32); the spectrograms show clear resolution of dynamic, rising spectral bands corresponding to the harmonics of the fundamental frequency. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

of temporal regularity (Pc–Nc). Furthermore, an asymmetry favoring the right hemispheric (RH) was observed only for the prototypical tonal variant (T2_250) suggesting the emergence of early stimulus-dependent hemispheric asymmetry (see Figure 13). The stimulus with the most gradual change in pitch acceleration, T2_250, appears to cross that threshold that leads to a RH asymmetry. Yet within the left hemisphere (LH), T2_250 was indistinguishable from T2_200. By assuming that the LH is preferentially engaged for mediating the categorical status of pitch, this finding suggests that T2_200 as well as T2_250 are better candidates as representations of Mandarin Tone 2. Thus, within this early cortical time window we begin to see the emergence of hemispheric preference, and how the complementary roles of the two hemispheres reflect influences from sensory and cognitive properties of the stimulus.

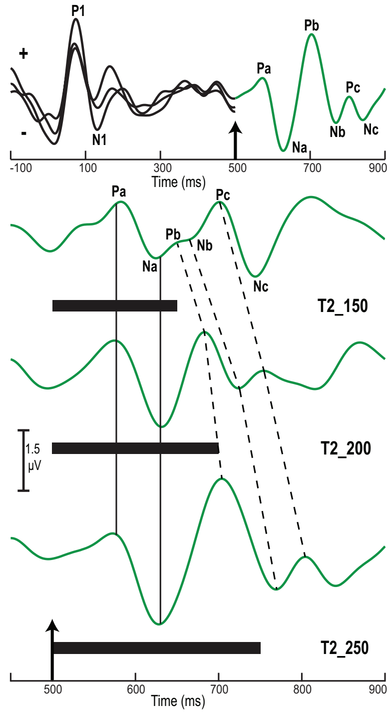


Figure 8. Grand averaged cortical evoked response components at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (green) are displayed in the top panel. The up arrow at 500 ms marks the boundary between the noise precursor segment and the pitch-eliciting stimulus. Na, Pb, and Nb are the most robust response components. CPR waveforms (green) elicited by the three stimuli (T2_150, T2_200, T2_250) are shown in the bottom panel. The up arrow at 500 ms marks the beginning of the pitch-eliciting stimulus. Solid black horizontal bars indicate the duration of each stimulus. Whereas Pa and Na do not change across stimuli (solid vertical lines), Pb, Nb, Pc, and Nc all show a systematic increase in peak latency (dashed vertical lines). Response amplitude for Na, Pb, and Nb increases from T2_150 to T2_250 in conjunction with decreasing pitch acceleration and increasing duration across stimuli. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

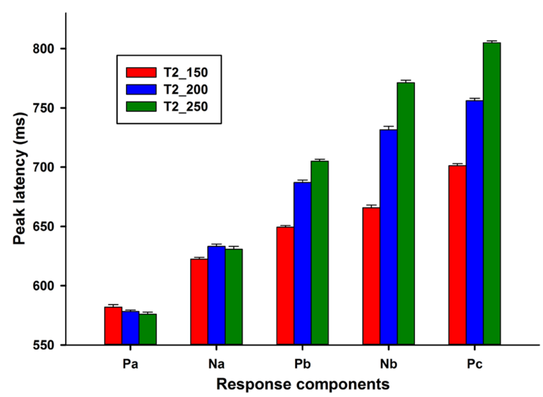


Figure 9. Mean peak latency of Fz response components as a function of stimulus. From stimulus onset (Pa) to stimulus offset (Pc), peak latencies show increasing differentiation across stimuli (T2_150; T2_200; T2_250) reflecting sensitivity to increases in pitch acceleration and duration. No changes in peak latency are evident for Pa in contrast to marked changes across stimuli for Pb, Nb, and Pc. Na (pitch onset) reveals that the peak latency of T2_150 is shorter than either T2_200 or T2_250. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

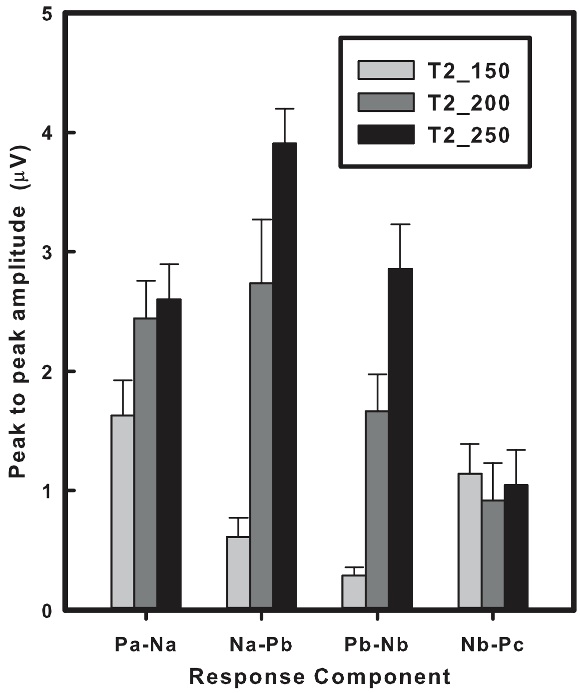


Figure 10. Mean peak-to-peak amplitude of Fz response components as a function of stimulus. Two pitch-relevant components (Pa-Na, Na–Pb, Pb–Nb) show that peak-to-peak amplitude increases steadily across the three stimuli (T2_150; T2_200; T2_250). In the case of Pa–Na, T2_250 is marginally higher in peak-to-peak amplitude than T2_150. Nb–Pc shows no differences in peak-to-peak amplitude between stimuli. Error bars.+/-1SE. **Reprinted from Krishnan et al., 2014 with permission from Elsevier R.V.**

Evidence for experience-dependent plasticity in the auditory cortex

Having described the basic characteristics of CPR components, we next examined how language experience (Mandarin vs. English) shapes the processing of different temporal attributes of pitch reflected in those CPR components using the same three variants of Tone 2 [64]. Results showed that the magnitude of CPR components (Na-Pb and Pb-Nb) and the correlation between these two components and pitch acceleration were stronger for the Chinese listeners (see Figures 11 and 12) compared to English listeners for stimuli that fell within the range of Tone 2 (acceleration rates presented in the pitch contours of T2_250 and T2_200 do occur in the language experience whereas the acceleration rate for the pitch contour T2_150 is well beyond rates in their language experience). Discriminant function analysis is used to determine which variables discriminate between two or more naturally occurring groups. In our application, it revealed that the Na-Pb component was more than twice as important as Pb-Nb in grouping listeners by language affiliation. This language-dependent effect suggests an experience-dependent increase in sensitivity to rapidly changing portions of pitch contours that occur in the native listeners' experience. Because enhanced sensitivity to time-varying dimensions (e.g., acceleration) is already present at the level of the brainstem [12, 27] it seems plausible that cortical pitch mechanisms may be reflecting, at least in part, this enhanced pitch input from the brainstem. We further note that experience-dependent enhancement of pitch was reflected primarily in the amplitude, instead of the latency, of CPR components. The more robust amplitude suggests greater temporal synchronization and improved synaptic efficiency of pitch-relevant neural activity among cortical neurons generating these CPR components. In contrast, absolute and interpeak latency may simply serve as discrete temporal event markers of neural activity indexing the temporal course of a pitch contour. That is, while amplitude may reflect the strength of the pitch relevant neural activity, the latencies may define the tine window in which this activity is occurring.

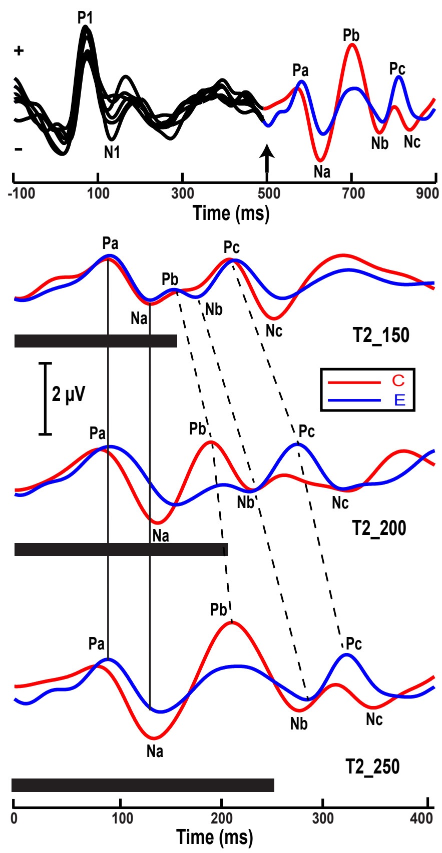


Figure 11. Grand averaged cortical evoked response components recorded at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (Chinese, red; English, blue) are displayed in the top panel. The up arrow at 500 ms marks the onset of the pitch-eliciting segment of the stimulus. Na, Pb, and Nb are the most robust response components. CPR waveforms elicited by the three stimuli (T2_150, T2_200, T2_250) are shown in the bottom panels. Solid black horizontal bars indicate the duration of each stimulus. Whereas Pa and Na do not change appreciably across stimuli (solid vertical lines), Pb, Nb, and Pc all show a systematic increase in peak latency (dashed vertical lines). Response amplitude for Na, Pb, and Nb increases from T2_150 to T2_250 in conjunction with decreasing pitch acceleration and increasing duration across stimuli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

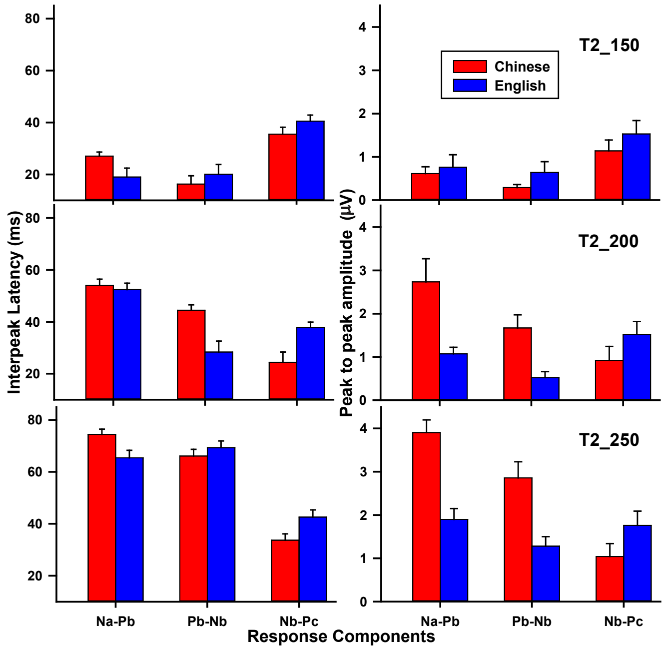


Figure 12. Mean interpeak latency (left panel) and peak-to-peak amplitude (right panel) of Na–Pb, Pb–Nb, and Nb–Pc components recorded at the Fz electrode site from T2_150 (top panel) to T2_250 (bottom panel) in both Chinese and English groups. Interpeak latencies increase across stimuli for Na–Pb and Pb–Nb in both groups. In the case of T2_200 (middle panel), Chinese interpeak latency is longer than English for Pb–Nb, but shorter than English for Nb–Pc. The Chinese group exhibits larger amplitude than the English group for Na–Pb and Pb–Nb in T2_200 and T2_250. Error bars = ±1 SE. **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

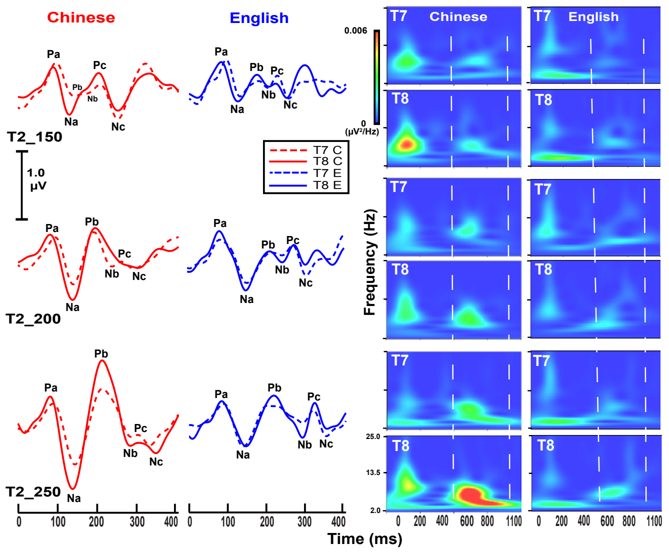


Figure 13. Grand average waveforms (two left panels) and their corresponding spectrograms (two right panels) of the CPR components for the two groups (Chinese in red and English in blue) recorded at electrode sites T7 (left temporal) and T8 (right temporal) for each of the three stimuli. Pitch-relevant CPR response components are generally greater in magnitude, and in particular for stimulus T2_250, for the Chinese group compared to the English group. The spectrograms in the right two panels show a large rightward asymmetry in the pitch-relevant neural activity (time window indicated by the dashed vertical lines) for the Chinese group only for stimulus T2_250. Note that while the waveform data time axis is referenced to the onset of the pitch segment (0-400 ms), the spectrogram is for the entire stimulus. The two vertical dashed lines in the spectrogram indicate the time window corresponding to CPR components. **Reprinted from Krishnan et al., 2014 with permission from Elsevier, R.V.**

In addition, a stronger stimulus-dependent, RH preference was observed for the Chinese group (see Figure 13). One view is that auditory processing occurs symmetrically in the core (primary auditory cortex), but asymmetrically in auditory-related areas beyond the core [65, 66]. We hypothesize that the language-dependent RH preference of the CPR response to T2_250 is due to an interaction with pitch-specific areas beyond the core that, in turn, are connected to higher order memory areas related to language. As such, it is an example of the essential interaction between sensory and cognitive components in pitch processing within the language domain in right auditory-related cortex. These findings are consistent with earlier ERP data showing emergence of experience-dependent hemispheric preference at early cortical levels of processing. For example, pure tones produced larger early cortical components (N19 mP30m) with a RH preference in musicians only; non-musicians show no hemispheric preference [67]. In response to musical stimuli, the N1 (change-N1, 100ms latency) component, related to pitch transition, showed greater amplitude at the right temporal electrode site in trained musicians [68]. Both these studies, using musically relevant stimuli point to experience-dependent enhancement of pitch processing in the right auditory cortex related to the music domain. We must also point out that our stimuli exhibit dynamic, curvilinear F0 trajectories that are representative of a Mandarin lexical tone. Interestingly, MEG recordings fail to observe any hemispheric differences with regard to either latency or amplitude of the pitch-relevant cortical component elicited by stimuli with flat pitch [29, 52, 69, 70]. This disparity in hemispheric asymmetry between dynamic and flat pitch patterns, seen in our Chinese listeners, further emphasizes both, the specialization of neural networks driven by behavioral relevance of sound and the importance of using ecologically-relevant stimuli to study pitch processing in the language domain.

Taken together, these findings suggest that long term language experience shapes early sensory level processing of pitch in the auditory cortex, and that the sensitivity of the CPR and hemispheric preference for processing pitch information may vary depending on the relative linguistic importance of specific temporal attributes of dynamic pitch. Our findings are consistent with earlier cross-language studies that have revealed experience-dependent neural plasticity at both subcortical and cortical levels of the brain [9, 12, 17, 19]. We believe that long-term experience shapes pitch mechanisms at early sensory levels of pitch processing in the auditory cortex. Like in the brainstem, they sharpen response properties of neural elements to enable optimal representation of linguistically relevant temporal attributes of native pitch contours.

NEURAL MECHANISM(S) FOR EARLY SENSORY LEVEL PITCH PROCESSING IN THE BRAINSTEM AND AUDITORY CORTEX

It is generally agreed that lateral Heschl's gyrus is the putative source for the pitch onset component (Na). Specific generator sources for the remaining pitch-relevant components (Pb, Nb, Pc, Nc) are yet to be determined. We speculate that these later components (Na-Pb, Pb-Nb) reflect neural activity from spatially distinct generators that represent later stages of sensory processing, relative to Na, along a pitch processing hierarchy. Whether pitch-relevant information extracted by these neural generators is based on a spectral and/or temporal code is unclear. At subcortical levels up to the midbrain, physiologic and computational modeling data support the possibility of either a purely temporal mechanism or a hybrid mechanism using both spectral and temporal information [2, 71-73]. There is evidence that neurons in primary auditory cortex exhibit temporal and spectral response properties that could enable these pitch-encoding schemes [74, 75], but it is not known whether they form a network with pitch-selective neurons to carry out this process. Unlike the subcortical auditory structures where periodicity and pitch are often represented by regular temporal patterns of action potentials that are phase-locked to the sound waveform, the most commonly observed code for periodicity and pitch within cortical neurons is a modulation of spike rates as a function of F0. It is possible that the wider temporal integration window at the cortical level may render the auditory cortical neurons too sluggish to provide phase-locked representations of periodicity within the pitch range [76]. Thus, it is not yet clear how cortical neurons transform the autocorrelation-like temporal analysis in the brainstem to a spike rate code to extract pitch-relevant information. One possibility is that the temporal code is transformed in to a response synchrony code where temporally coherent activity from the subcortical stages will produce greater spike rate, yielding larger response amplitude at the cortical level. Interestingly, Micheyl, Strater & Oxenham [77] show (based on analyses of statistical properties of the spike rates of virtual neural units, that have similar frequency tuning and spike rate characteristics as auditory cortical neural units) that sufficient statistical information is present in the population spike rate to account for small differences in frequency (pitch) and intensity (loudness).

NEURAL MECHANISMS MEDIATING EXPERIENCE-DEPENDENT PLASTICITY FOR PROCESSING PITCH RELEVANT INFORMATION IN THE BRAINSTEM AND AUDITORY CORTEX

Experience-dependent enhancement of pitch representation for Chinese listeners most likely reflects an interaction between higher-level cognitive processes and early sensory-level processing to improve representations of behaviorally-relevant features that contribute optimally to perception. It is our view that long-term experience shapes this adaptive process wherein the top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural responses to specific behaviorally-relevant attributes of the stimulus. The goal clearly is to achieve optimal correspondence between the sensory representations and the resulting percept at all levels of processing [78]. Evidence for this signal selectivity, mediated through top-down influence, comes from response properties of cortical neurons in animal models, that show a selective increase in responsiveness and shifts in best frequencies toward task-relevant, target stimuli [79-81]; and selective expansion of receptive fields for stimulus features that are being learned [82]. In the case of humans, the top-down influence mediated by the corticofugal system likely shapes the enhancement of brainstem pitch representation resulting from short-term auditory training [83, 84]; long-term linguistic experience [12, 27, 35]; and musical training [48-50, 85-87]. The reverse hierarchy theory (RHT) provides a representational hierarchy to describe the interaction between sensory input and top-down processes to guide plasticity in primary sensory areas [88, 89]. This theory suggests that neural circuitry mediating a certain percept can be modified starting at the highest representational level and progressing to lower levels in search of more refined high resolution information to optimize percept. The RHT has been invoked as a plausible explanation for topdown influences on cortical [90] and subcortical sensory processing [38, 91]. Consistent with this theory, it is possible that sensory-level representation of spectrotemporal features related to pitch in the brainstem is more precise than the more labile, spatiotemporally broader, pitch-relevant information in the auditory cortex [93-95]. Indeed, fine-grained, spectrotemporal details that are present in the sustained brainstem response are absent in transient, cortical pitch response components. We nevertheless observe a close correspondence between cortical and brainstem responses when manipulating the degree of pitch salience [32].

One proposed circuitry, mediating learning-induced plasticity; the colliculo-thalamo-cortico-collicular loop [96] may be invoked to mediate the RHT described above. This circuitry is comprised of bottom-up (colliculo-thalamic and thalamo-cortical) and top-down (corticofugal) projections that form a tonotopic loop. It incorporates several neuromodulatory inputs that form a core neural circuit mediating sound-specific plasticity associated with perceptual learning. Auditory stimuli and neuromodulatory inputs are believed to induce large-scale, frequency-specific plasticity in the loop. It is also possible that bottom-up as well as local top-down cortical inputs may jointly influence pitch processing as reflected in the CPR components. In the case of the former, enhanced representations from brainstem pitch mechanisms are functionally reorganized by top-down influence during the critical period of language acquisition. As a result, brainstem responses constitute an indirect reflection of inputs from the corticofugal system. Once this reorganization is complete, local mechanisms in the brainstem and auditory cortex would be sufficiently robust to extract linguistically-relevant pitch information optimally without an engaged, online corticofugal influence [97]. Indeed, the strong correlation between neural representations relevant to pitch salience at the brainstem and early cortical levels of processing suggests that sensory processing at the brainstem level may be driving early preattentive sensory processing relevant to pitch at the cortical level [32]. In the case of humans, top-down processes likely shape the reorganization of the sensory processing of pitch-relevant information in the brainstem and early sensory level processing in the auditory cortex to enhance pitch extraction in earlier stages of language development when adaptive plasticity presumably would be most vigorous [98, 99]. The slower time constants of corticofugal processing render it much too sluggish to effectively influence a dynamic pitch pattern over its entire duration [100]. Nonetheless, its adaptive properties would still be able to facilitate extraction of behaviorally-relevant information under degraded listening conditions and during training protocols.

CONCLUSIONS

While language experience shapes pitch processing at both subcortical and cortical levels, neural representations are transformed and fundamentally different at each biological level of abstraction. The representation of pitch relevant information in the brainstem is more fine-grained spectrotemporally as it reflects sustained neural phase-locking to pitch relevant periodicities contained in the dynamic stimulus. In contrast, the cortical representation is coarser. That is, the cortical pitch relevant neural activity reflects primarily a series of distinct transient temporal neural events marking only certain temporal attributes of the pitch contour. These differences notwithstanding, we believe that long-term language experience shapes adaptive, hierarchical pitch processing. Top-down connections provide selective gating of inputs to both cortical and subcortical structures to enhance neural representation of behaviorally-relevant attributes of the stimulus and instantiate local mechanisms that exhibit enhanced representation of behaviorally-relevant pitch attributes. The ability to record brainstem and cortical pitch relevant responses concurrently may also provide a new window to evaluate the online interplay between feedforward and feedback components in the processing of pitch-relevant information at the level of the brainstem and the auditory cortex, and how experience shapes pitch processing at cortical and subcortical levels. The selection of dynamic pitch contours representative of lexical tone will enable us to evaluate the influence of language experience on the latency and amplitude of these cortical pitch response components in subsequent experiments. The challenge is to develop experiments that systematically manipulate pitch attributes in order to optimally evaluate the relationship between representation of pitch relevant information at the brainstem and cortical levels. The results of these experiments would be essential to develop a framework to understand both the nature of interplay between levels of processing and, interactions between sensory and cognitive processes influencing pitch representation. Complementary studies using MEG will be crucial to determine the anatomical sources of these components in an effort to shed more light on specific cortical generators contributing to the hierarchical stages of pitch processing, and how experience may shape these processes.

ACKNOWLEDGEMENTS

Research supported by NIH 5R01DC008549-06 (A.K.). Thanks to Jingyi Zhu and Longjie Cheng for their assistance with statistical analysis (Department of Statistics); JillianWendel and Chandan Hunsur Suresh for their help with data acquisition and graphics, respectively. Reprint requests should be addressed to Ananthanarayan Krishnan, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907-2038,USA, or via email: rkrish@purdue.edu.

REFERENCES

[1] Oxenham, A. J. (2012). Pitch perception. *Journal of Neuroscience*, **32**(39), 13335-13338. doi:10.1523/JNEUROSCI.3815-12.2012

[2] Plack, C. J., Oxenham, A. J., & Fay, R. R. (Eds.). (2005). *Pitch: neural coding and perception (Vol. 24)*. New York: Springer

[3] Kumar, S., Stephan, K.E., Warren, J.D., Friston, K.J., Griffiths, T.D. (2007). Hierarchical processing of auditory objects in humans. *PLoS Computational Biology* **3**: e100. doi:10.1371/journal.pcbi. 0030100.

[4] Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, 4(6), 633-637.

[5] Griffiths, T.D., Buchel, C., Frackowiak, R.S., & Patterson, R.D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, **1**(5), 422–427.

[6] Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., & Griffiths, T.D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, **36**(4), 767–776.

[7] Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2007). The effect of temporal context on the sustained pitch response in human auditory cortex. *Cerebral Cortex*, **17**(3), 552–561, http://dx.doi.org/10.1093/cercor/bhj180

[8] Besson, M., Chobert, J., & Marie, C. (2011). Language and music in the musician brain. *Language and Linguistics Compass*, **5**(9), 617–634, http://dx.doi.org/10.1111/j.1749-818x.2011.00302

[9] Gandour, J. T., & Krishnan, A. (2014). Neural bases of lexical tone. In: *H. Winskel, & P. Padakannaya (Eds.), Handbook of South and Southeast Asian psycholinguistics* (pp. 339–349). Cambridge, UK: Cambridge University Press.

[10] Koelsch, S. (2012). *Brain & Music*. Chichester, UK: Wiley-Blackwell.

[11] Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, **16**(2), 105–110.

[12] Krishnan, A., Gandour, J.T., & Bidelman, G.M. (2012). Experience-dependent plasticity in pitch encoding: From brainstem to auditory cortex. *Neuroreport*, **23**(8), 498–502.

[13] Meyer, M. (2008). Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitshcrift fur Neuropsycholgie*, **19**(2), 101–115.

[14] Munte, T. F., Altenmuller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews: Neuroscience*, **3**(6), 473–478, http://dx.doi.org/10.1038/nrn843.

[15] Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, **11**(9), 369–372.

[16] Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *The European Journal of Neuroscience*, **30**(8), 1636–1642, http://dx.doi.org/10.1111/j.1460-9568.2009.06955.x.

[17] Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, **10**(7), e1001372. doi:10.1371/journal.pbio.1001372

[18] Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, **6**(1), 37–46.

[19] Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, **363**(1493), 1087-1104. doi: J412P80575385013 [pii] 10.1098/rstb.2007.2161.

[20] Chandrasekaran, B., & Kraus, N. (2010). *The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology*, **47**(2), 236-246. doi: PSYP928 [pii] 10.1111/j.1469-8986.2009.00928.x

[21] Maddieson, I. (1978). Universals of tone. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 2, pp. 335-365). Stanford, CA: Stanford University Press.

[22] Yip, M. (2003). Tone. New York: Cambridge University Press

[23] Abramson, A. S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments*. Bloomington: Indiana U. Research Center in Anthropology, Folklore, and Linguistics, Pub. 20.

[24] Gandour, JT. (1994). Phonetics of tone. In: Asher, R.; Simpson, J., editors. *The encyclopedia of language & linguistics*. New York: Pergamon Press; p. 3116-3123.

[25] Xu, Y. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, **33**, 319-337.

[26] Griffiths, T. D., Warren, J. D., Scott, S. K., Nelken, I., & King, A. J. (2004). Cortical processing of complex sound: A way forward? *Trends in Neurosciences*, **27**(4), 181–185.

[27] Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, **110**(3), 135-148. doi: S0093-934X(09)00042-X [pii] 10.1016/j.bandl.2009.03.005

[28] Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C.,& Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, **13**(7), 765–772.

[29] Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). Evidence of pitch processing in the N100m component of the auditory evoked field. *Hearing Research*, **213**(1-2), 88–98.

[30] Griffiths, T.D., Kumar, S., Sedley,W., Nourski, K.V., Kawasaki, H.,Oya, H.,et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, **20**(12),1128–1132.

[31] Schonwiesner, M., & Zatorre, R.J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, **187**(1), 97–105.

[32] Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., & Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, **50**(12), 2849-2859. doi: 10.1016/j.neuropsychologia.2012.08.013

[33] Krishnan, A. (2007). Human frequency following response. In R.F. Burkard, M. Don & J. J. Eggermont (Eds.), *Auditory evoked potentials: Basic principles and clinical application* (pp. 313-335). Baltimore: Lippincott Williams & Wilkins.

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

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

[36] Krishnan, A. (2002). Human frequency following response: Representation of steady-state vowels. *Hearing Research*, **166**,192-201.

[37] Krishnan, A., & Plack, C.J. (2011). Neural encoding in the human brainstem relevant to the pitch of complex tones. *Hearing Research*, **275**:110–119.

[38] Krishnan, A., Bidelman, G. M., & Gandour, J. T. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, **268**(1-2), 60-66. doi: 10.1016/j.heares.2010.04.016

[39] Krishnan A., Gandour, J.T., Bidelman, G.M., & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, **20**:408–413.

[40] Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, **21**(6), 1092-1105. doi: 10.1162/jocn.2009.21077

[41] Swaminathan, J., Krishnan, A., & Gandour, J.T.(2008). Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*, **19**:1163–1167.

[42] Bidelman, G.M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, **22**:212–216.

[43] Gockel, H.E., Carlyon, R.P., Mehta, A., and Plack, C.J. (2011). The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *Journal of the American Association for Research in Otolaryngology*, **12**,767-782

[44] Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010b). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, **23**(1), 81-95. doi: 10.1016/j.jneuroling.2009.09.001

[45] Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010a). Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research*, **1313**, 124-133. doi10.1016/j.brainres.2009.11.061

[46] Krishnan, A., Gandour, J. T., Smalt, C. J., & Bidelman, G. M. (2010). Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain and Language*, **114**(3), 193-198. doi: 10.1016/j.bandl.2010.05.004

[47] Xu, Y., Krishnan, A., & Gandour, J.T. (2006). Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport*, **17**:1601–1605.

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

[49] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*. **77**:1–10.

[50] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011c). Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neuroscience Letters*. **503**:203–207.

[51] Alku, P., Sivonen, P., Palomaki, K., & Tiitinen, H. (2001). The periodic structure of vowel sounds is reflected in human electromagnetic brain responses. *Neuroscience Letters*, **298**(1), 25–28.

[52] Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *Neuroimage*, **22**, 755-766.
[53] Lutkenhoner, B., Seither-Preisler, A., & Seither, S. (2006). Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. *Neuroimage*, **30**(3), 927–937.
[54] Soeta, Y., & Nakagawa, S. (2008). The effects of pitch and pitch strength on an auditory evoked N1m. *Neuroreport*, **19**(7), 783–787.
[55] Yrttiaho, S., Tiitinen, H., May, P.J., Leino, S., & Alku, P. (2008). Cortical sensitivity to periodicity of speech sounds. *Journal of the Acoustical Society of America*, **123**(4), 2191–2199.
[56] Chait, M., Poeppel, D., & Simon, J. Z. (2006). Neural response correlates of detection of monaurally and binaurally created pitches in humans. *Cerebral Cortex* (New York, NY: 1991), **16**(6), 835–848.
[57] Ritter, S., GunterDosch, H., Specht, H. J., & Rupp, A.(2005). Neuromagnetic responses reflect the temporal pitch change of regular interval sounds. *Neuroimage*, **27**(3), 533–543.
[58] Gutschalk, A., Patterson, R.D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage*, **15**(1), 207–216.
[59] Griffiths, T.D., Kumar, S., Sedley,W., Nourski, K.V., Kawasaki, H.,Oya, H.,et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, **20**(12), 1128–1132.
[60] Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, **436**(7054), 1161–1165.
[61] Johnsrude, I.S., Penhune,V.B., & Zatorre, R.J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, **123**(Part1), 155–163.
[62] Penagos, H., Melcher, J.R., & Oxenham, A.J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, **24**(30), 6810–6815.
[63] Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2014a). Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*, 59C, 1-12.
[64] Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2015). Language experience enhances early cortical pitch-dependent responses. *Journal of Neurolinguistics*, **33**, ,128-148.
[65] Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, **41**(1), 245–255.
[66] Poeppel, D., Idsardi, W.J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, **363**(1493).
[67] Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A. & Rupp, A. (2002) Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, **5**, 688–694.
[68] Itoh, K., Okumiya-Kanke, Y., Nakayama, Y., Kwee, I.L., & Nakada,T. (2012). Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *European Journal of Neuroscience*, **36**(11), 3580–3592.
[69] Hari, R., Pelizzone, M., Makela, J.P., Hallstrom, J., Leinonen, L. & Lounasmaa, O.V. (1987). Neuromagnetic responses of the human auditory cortex to on-and offsets of noise bursts. *Audiology*, **26**(1), 31–43.
[70] Lutkenhoner, B., & Steinstrater, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neuro-Otology*, **3**, 191–213.
[71] Cariani, P.A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I.Pitch and pitch salience. *Journal of Neurophysiology*, **76**(3), 1698–1716.
[72] Cariani, P.A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *Journal of Neurophysiology*, **76**(3), 1717–1734.
[73] Cedolin, L., & Delgutte, B. (2005). Pitch of complex tones: Rate-place and interspike interval representations in the auditory nerve. *Journal of Neurophysiology*, **94**(1), 347–362.
[74] Lu, T., Liang, L., & Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience*, **4**(11), 1131-1138. doi:10.1038/nn737
[75] Steinschneider, M., Reser, D. H., Fishman, Y. I., Schroeder, C. E., & Arezzo, J. C. (1998). Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *Journal of the Acoustical Society of America*, **104**(5), 2935-2955.
[76] Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Cortical encoding of pitch: recent results and open questions. *Hearing Research*, **271**(1-2), 74-87. doi: 10.1016/j.heares.2010.04.015
[77] Micheyl, C., Schrater, P.R., & Oxenham, A.J. (2013). Auditory frequency and intensity discrimination explained using a cortical population rate code. *PLoS Computational Biology*, **9**(11):e1003336. doi: 10, 1371/journal.pcbi.1003336.
[78] Gilbert, C. D., & Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, **54**(5), 677-696. doi: 10.1016/j.neuron.2007.05.019
[79] Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, **6**(11), 1216-1223. doi: 10.1038/nn1141
[80] Lee, C. C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, **14**(1), 108-114. doi: 10.1038/nn.2713
[81] Weinberger, N. M. (2011). Reconceptualizing the primary auditory cortex: learning, memory and specific plasticity. In J. A. Winer & C. E. Schreiner (Eds.), *The auditory cortex* (pp. 465-491). New York: Springer
[82] Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *Journal of Neuroscience*, **26**(18), 4970-4982.
[83] Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, **156**(1), 95-103.
[84] Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, **20**, 1892-1902.
[85] Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, **29**(42), 13165-13171.
[86] Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(40), 15894-15898. doi: 0701498104 [pii] 10.1073/pnas.0701498104

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

[88] Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, **8**(10), 457-464. doi: 10.1016/j.tics.2004.08.01 1S1364-6613(04)00215-3 [pii]

[89] Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: the case of speech in noise. *PLoS Biology*, **6**(5), e126. doi: 07-PLBI-RA-2244 [pii]10.1371/journal .pbio.0060126

[90] Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, **128**(1), 34-40. doi: 10.1016/j. bandl.2013.11.006

[91] Banai, K., Abrams, D., & Kraus, N. (2007). Sensory-based learning disability: Insights from brainstem processing of speech sounds. *International Journal of Audiology*, **46**(9), 524-532. doi: 781872134 [pii] 10.1080/

[92] Chechik, G., Anderson, M. J., Bar-Yosef, O., Young, E. D., Tishby, N., & Nelken, I. (2006). Reduction of information redundancy in the ascending auditory pathway. *Neuron*, **51**(3), 359-368. doi: 10.1016/j.neuron.2006.06.030

[93] Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*, **23**(13), 5799-5804.

[94] Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends in Neuroscience*, **28**(5), 255-263 doi: 10.1016/ j.tins. 2005.03.009

[95] Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, **11**(10), 946-953.

[96] Xiong, Y., Zhang, Y., & Yan, J. (2009). The neurobiology of sound-specific auditory plasticity: A core neural circuit. *Neuroscience and Biobehavioral Reviews*, **33**(8), 1178-1184. doi: doi:10.1016/j.neubiorev.2008.10.006

[97] Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, **13**(2), 253-260. doi: 10.1038/nn.2466

[98] Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in Neurobiology*, **82**(3), 109-121. doi: S0301-0082(07)00073-1[pii] 10.1016/j.pneurobio.2007.03.005

[99] Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews*, **56**(1), 259-269. doi: S0165-0173(07)00187-7 [pii]10.1016/j. brainresrev.2007.07.021

[100] Dean, I., Robinson, B. L., Harper, N. S., & McAlpine, D. (2008). Rapid neural adaptation to sound level statistics. *Journal of Neuroscience*, **28**(25), 6430-6438. doi: 28/25/6430 [pii] 10.1523/JNEUROSCI.0470-08.2008