



Published in final edited form as:

*J Neurolinguistics*. 2017 February ; 41: 38–49. doi:10.1016/j.jneuroling.2016.09.005.

## Language-dependent changes in pitch-relevant neural activity in the auditory cortex reflect differential weighting of temporal attributes of pitch contours

Ananthanarayan Krishnan<sup>a</sup>, Jackson T. Gandour<sup>a</sup>, Yi Xu<sup>b</sup>, and Chandan H. Suresh<sup>a</sup>

<sup>a</sup>Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN USA

<sup>b</sup>Department of Speech, Hearing and Phonetic Sciences, University College London, UK

### Abstract

There remains a gap in our knowledge base about neural representation of pitch attributes that occur between onset and offset of dynamic, curvilinear pitch contours. The aim is to evaluate how language experience shapes processing of pitch contours as reflected in the amplitude of cortical pitch-specific response components. Responses were elicited from three nonspeech, bidirectional (falling-rising) pitch contours representative of Mandarin Tone 2 varying in location of the turning point with fixed onset and offset. At the frontocentral Fz electrode site, Na–Pb and Pb–Nb amplitude of the Chinese group was larger than the English group for pitch contours exhibiting later location of the turning point relative to the one with the earliest location. Chinese listeners' amplitude was also greater than that of English in response to those same pitch contours with later turning points. At lateral temporal sites (T7/T8), Na–Pb amplitude was larger in Chinese listeners relative to English over the right temporal site. In addition, Pb–Nb amplitude of the Chinese group showed a rightward asymmetry. The pitch contour with its turning point located about halfway of total duration evoked a rightward asymmetry regardless of group. These findings suggest that neural mechanisms processing pitch in the right auditory cortex reflect experience-dependent modulation of sensitivity to weighted integration of changes in acceleration rates of rising and falling sections and the location of the turning point.

### Keywords

auditory; pitch encoding; iterated rippled noise; cortical pitch response; experience-dependent neuroplasticity; Mandarin

---

Corresponding author: Jackson T. Gandour, PhD, Purdue University, Department of Speech Language Hearing Sciences, Lyles-Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907 USA, gandour@purdue.edu, PHONE: 1-765-237-1366. Correspondence to: A. Krishnan, Purdue University, Department of Speech Language Hearing Sciences, Lyles Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907-2122, USA, rkrish@purdue.edu (A. Krishnan).

**Postal address:** Lyles Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907-2122

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

## 1. Introduction

Pitch is an important information-bearing perceptual attribute that provides an excellent window for studying language-dependent effects on pitch processing at both subcortical and cortical levels. Tone languages are especially advantageous for investigating the linguistic use of pitch because variations in pitch patterns at the syllable level may be lexically significant (Laver, 1994, p. 465). Neural representation of pitch may be influenced by one's experience with language (or music) at subcortical as well as cortical levels of processing (for reviews, see Gandour & Krishnan, 2016; Johnson, Nicol, & Kraus, 2005; Kraus & Banai, 2007; Krishnan & Gandour, 2009, 2014; Patel & Iversen, 2007; Tzounopoulos & Kraus, 2009; Zatorre & Baum, 2012; Zatorre & Gandour, 2008). Of special interest herein, Jia, Tsang, Huang, and Chen (2015) have recently demonstrated—using the mismatch negativity (MMN)—that processing of lexical tone relies on both acoustic and linguistic factors at an early cortical stage of processing.

Experience-dependent enhancement of pitch relevant phase-locked neural activity in the auditory brainstem has been observed consistently for rapidly-changing pitch sections that occur after the turning point of high rising Mandarin Tone 2 (T2) in nonspeech contexts (Bidelman, Gandour, & Krishnan, 2011; Krishnan, Swaminathan, & Gandour, 2009). In response to a continuum of nonspeech variants of T2 varying in acceleration rate (Krishnan, Gandour, Smalt, & Bidelman, 2010), both Mandarin- and English-speaking groups show decreasing pitch strength of brainstem responses with increasing acceleration rates. Yet Chinese are more resistant to degradation even in response to scaled variants of T2 with higher acceleration rates that fall proximal to or outside the normal voice pitch range. Across studies of T2 at the level of the auditory brainstem, the location of the turning point was a constant.

At the level of the cerebral cortex, the EEG-derived, human cortical pitch response (CPR) transient components (Na–Pb, Pb–Nb) reveal that Mandarin-speaking natives are sensitive to changes in pitch acceleration as elicited by three within-category, nonspeech variants of T2 (Krishnan, Gandour, Ananthakrishnan, & Vijayaraghavan, 2014). These stimuli varied both in duration and location of turning point. 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). Pc–Nc marks unambiguously the stimulus offset. A strong correlation with pitch acceleration is observed for Na–Pb and Pb–Nb, putative indices of pitch-relevant neural activity associated with the more rapidly-changing sections of the pitch contour. A crosslanguage study shows that the magnitude of CPR components Na–Pb and Pb–Nb is larger for Chinese than English listeners in response to all three variants of T2 (Krishnan, Gandour, Ananthakrishnan, & Vijayaraghavan, 2015). Using the same set of pitch stimuli as in Krishnan, et al. (2010), Chinese listeners show greater amplitude than English for both Na–Pb and Pb–Nb at frontocentral (Fz) and temporal (T7/T8) electrode sites in response to the two pitch contours with acceleration rates that fall inside the normal voice pitch range, i.e., those that fall within the bounds of one's native language (Krishnan, Gandour, & Suresh, 2015a). As indexed by Na–Pb and Pb–Nb amplitude at the temporal sites, a right-sided preference was observed for the Chinese group only. Amplitude of the Chinese group was greater than that of the English only over the right

temporal site. We infer that the neural mechanism(s) underlying processing of pitch in the right auditory cortex reflect experience-dependent modulation of sensitivity to changes in acceleration rates on the section of the pitch contour from turning point to offset in T2.

To extend this line of research on sensitivity of cortical responses to changes in the temporal dynamics throughout an entire pitch contour, we now recognize the need to design stimulus sets that include the following variables: onset, offset, and turning point. A pitch contour necessarily involves a movement from its onset to offset. What happens between onset and offset can best be expressed acoustically as changes in acceleration (or velocity). In natural speech production, this movement is dynamic, and curvilinear in nature. Using Cantonese lexical tones with varying onsets and offsets (Tsang, Jia, Huang, & Chen, 2011), the size or latency of MMN did not vary depending upon the location of the turning point in pitch contours. The latency of P3a—ERP component associated with automatic switching of attention induced by unexpected change in stimulus event—was delayed to the same extent irrespective of the location of the turning point. To the best of our knowledge, there is no published report of a scalp-recorded, event-related brain potentials (ERP) study of pitch processing in which the turning point is systematically varied while holding onset, offset and duration constant (cf. behavioral; Moore & Jongman, 1997). In previous studies of Mandarin lexical tones, turning point covaried with duration (Krishnan, et al., 2014; Krishnan, Gandour, Ananthakrishnan, et al., 2015) or turning point was fixed along with onset and duration (Krishnan, Gandour, et al., 2015a), or location of turning points were invariant within stimulus sets (Bidelman & Lee, 2015).

As our knowledge grows about the CPR, it is important to establish the lower and upper bounds of its sensitivity to various pitch attributes (e.g., acceleration). We have therefore incorporated a quantitative model for generating voice fundamental frequency (F0) contours of speech (Prom-on, Xu, & Thipakorn, 2009). This model is based on biophysical and linguistic assumptions about the mechanisms of F0 production. It permits us to design stimulus sets that address specific research questions regarding the sensitivity of cortical and/or brainstem responses to changes in the temporal dynamics of a pitch contour. In this study, the stimulus set represents a continuum of bidirectional pitch contours that exhibit changes in acceleration in two sections of the pitch contour. The first section spans from onset to turning point; the second section, from turning point to offset. We are able to change acceleration systematically in both sections by holding F0 onset and offset constant, and progressively delay the temporal location of the turning point in the pitch contour. The result is a three-step, F0 turning point continuum in which pitch contours represent within-category variants of T2 that may be produced on isolated monosyllables.

The specific aim of this study is to determine how changes in the temporal location of the turning point in the bidirectional pitch contours influence amplitude of CPR components as a function of language experience (Mandarin, English). Stimulus comparisons at the frontocentral electrode site allow us to assess the effects of changes in location of the turning point. At the right temporal site, the pattern of changes in CPR components is expected to reflect temporally distinct, language-dependent differential weighting of sensory and extrasensory effects. More generally, we infer that enhancement of pitch-relevant neural activity reflects optimal, differentially weighted, integration of pitch relevant neural activity

to changes in acceleration rates of the falling and rising segments and the location of the turning point.

## 2. Methods

### 2.1. Participants

EEG data were recorded from a total of twelve native speakers of Mandarin Chinese (6 male, 6 female) and English (6 male, 6 female) recruited from the Purdue University student body. All exhibited normal hearing sensitivity at audiometric frequencies between 500 and 4000 Hz and reported no previous history of neurological or psychiatric illnesses. They were matched in age (Chinese:  $24.0 \pm 3.5$  years; English:  $21.7 \pm 1.7$ ), years of formal education (Chinese:  $17.7 \pm 2.6$  years; English:  $15.5 \pm 0.9$ ), and were strongly right handed (Chinese:  $91.0 \pm 15.3\%$ ; English:  $92.0 \pm 12.6\%$ ) as measured by the laterality index of the Edinburgh Handedness Inventory (Oldfield, 1971; Somers, et al., 2015). All Chinese participants were born and raised in mainland China. None had received formal instruction in English before the age of nine ( $11.5 \pm 1.5$  years). Self-ratings of their English language proficiency on a 7-point Likert-type scale ranging from 1 (very poor) to 7 (native-like) for speaking and listening abilities were, on average, 5 and 5.1, respectively (Li, Sepanski, & Zhao, 2006). Their daily usage of Mandarin and English, respectively, was reported to be 63.8% and 36.1%. As determined by a music history questionnaire (Wong & Perrachione, 2007), all Chinese and English participants had less than two years of musical training (Chinese,  $0.46 \pm 0.78$  years; English,  $1.0 \pm 1.2$ ) on any combination of instruments. No participant had any training within the past five years. Each participant was paid and gave informed consent in conformity with the 2013 World Medical Association Declaration of Helsinki and in compliance with a protocol approved by the Institutional Review Board of Purdue University.

### 2.2. Stimuli

Three nonspeech stimuli were designed to represent a falling-rising pitch continuum varying in the location of the turning point (TP). The location of the turning point of TP1–TP3 in order of increasing duration in ms and % of total duration was 74 (30%), 116 (46%), and 156 (62%). F0 onset and offset were fixed at 140 and 160 Hz, respectively. Because of fixed onset and offset, any change in location of the turning point resulted in a change in the falling and rising acceleration rates. In other words, rates of acceleration from onset to turning point and from turning point to offset were interdependent. For example, the longer the location of the turning point, the larger the acceleration rate to offset. TP stimuli were divided into two sections, respectively, in order to characterize the pitch acceleration rates of the falling and rising sections of the F0 contour (Fig. 1): F0 onset to turning point, falling (solid); turning point to F0 offset, rising (dash). Table 1 lists the duration values of the falling and rising sections per stimulus. Table 2 displays the mean velocity (st/s) of the falling and rising sections of each TP stimulus. In comparison to data on mean speed of pitch change (Xu & Sun, 2002, p. 1407, Table VII), TP1 and TP3 exceed the mean for the falling and rising section, respectively. TP2 does not exceed the mean value for either section. All three stimuli fall within the range of combinations of turning point and F0 from onset to turning point that trigger identification of T2 (Moore & Jongman, 1997). All

three TP stimuli are judged by the third author to be representative of T2 (cf. Xu, 1997). Across the TP continuum, the excursion size from onset to turning point was about 4 st; from turning point to offset, about 6 st (Supplementary online material: sound files of TP stimuli).

As a check on the biomechanical plausibility of the F<sub>0</sub> contours, we applied PENTAtainer, a computational modeling tool for estimating underlying control parameters of tone production (Prom-on, et al., 2009; Xu & Prom-on, 2010–2015). PENTAtainer extracted an optimal parameter set, i.e., one that can generate an F<sub>0</sub> contour that best fits each stimulus. Each parameter set consists of: target height, the ideal pitch register; target slope, the ideal velocity of F<sub>0</sub> movement; and target strength, the rate at which the two-dimensional target is approached. Results of the parameter estimation showed it was possible to find a set of parameters to generate a nicely fitted contour for all three TP stimuli, meaning that they were representative of possible pitch contours in natural speech.

Iterated rippled noise was used to create these stimuli by applying polynomial equations that generate dynamic, curvilinear pitch patterns (Swaminathan, Krishnan, & Gandour, 2008). IRN enables us to preserve dynamic variations in pitch of auditory stimuli that lack formant structure, temporal envelope, and recognizable timbre characteristic of speech. IRN stimuli were created by delaying Gaussian noise (80–4000Hz) and adding it back on itself in a recursive manner (Yost, 1996a). The pitch of IRN corresponds to the reciprocal of the delay (1/d); its salience grows with the number of iterations with little or no change in salience beyond an iteration step of 32 (Yost, 1996b), the upper limit used here. These three IRN stimuli were presented in a paradigm consisting of three segments: a 250 ms pitch segment (TP1–TP3) preceded by a 750 ms noise segment and followed by a 235 ms noise segment.

All stimuli were presented binaurally at 80 dB SPL through magnetically-shielded tubal insert earphones (ER-3A<sup>®</sup>; Etymotic Research, Elk Grove Village, IL, USA) with a fixed onset polarity (rarefaction) and a repetition rate of 0.56/s. The overall root-mean-square level of each segment was equated such that there was no discernible difference in intensity across the three segments. All stimuli were generated and played out using an auditory evoked potential system (SmartEP<sup>®</sup>, Intelligent Hearing Systems; Miami, FL, USA).

### 2.3. Cortical evoked response data acquisition

Participants reclined comfortably in an electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement to minimize myogenic artifacts, and to ignore the stimuli as they watched a silent video (minus subtitles) of their choice throughout the recording session. The EEG was acquired continuously (5000 Hz sampling rate; 0.3 to 2500 Hz analog band-pass) through the ASA-Lab EEG<sup>®</sup> system (ANT Inc., The Netherlands) using a 32-channel amplifier (REFA8-32, TMS International BV) and WaveGuard electrode cap (ANT Inc., The Netherlands) with 32-shielded sintered Ag/AgCl electrodes configured in the standard 10–20-montage. Because the primary objective of this study was to characterize cortical pitch components, the EEG acquisition electrode montage was limited to 9 electrode locations: Fpz, AFz, Fz, F3, F4, Cz, T7, T8, M1, M2. The AFz electrode served as the common ground, and the common average of all connected unipolar electrode inputs served as default

reference for the REFA8-32 amplifier. An additional bipolar channel with one electrode placed lateral to the outer canthi of the left eye and another electrode placed above the left eye was used to monitor artifacts introduced by ocular activity. Inter-electrode impedances were maintained below 10 k $\Omega$ . For each stimulus, EEGs were acquired in two blocks of 1000 sweeps each. Stimulus presentation order was randomized both within and across participants. The experimental protocol took about 2 hours to complete.

#### 2.4. Extraction of CPR

CPR responses were extracted off-line from the EEG files. To extract the CPR components, EEG files were first down sampled from 5000 Hz to 1024 Hz. They were then digitally band-pass filtered (2–25 Hz, Butterworth zero phase shift filter with 24 dB/octave rejection rate) to enhance the transient components and minimize the sustained component. Sweeps containing electrical activity exceeding  $\pm 50 \mu\text{V}$  were rejected automatically. Subsequently, averaging was performed on all 8 unipolar electrode locations using the common reference to allow comparison of CPR components at the right frontal (F4), left frontal (F3), right temporal (T8), and left temporal (T7) electrode sites to evaluate asymmetry effects. Given the poor spatial resolution of EEG even using multiple electrodes, the focus here is not to localize the source of the CPRs with just two electrodes, but to characterize the relative difference in the pitch-related neural activity over the widely separated left and right temporal electrode sites. In previous crosslanguage CPR studies, we have consistently observed robust differences in CPR neural activity over the T7 and T8 electrode sites that reflect a functional, experience-dependent rightward asymmetry (e.g., Krishnan, et al., 2014). The re-referenced electrode site, Fz linked (T7/T8), was used to characterize the transient pitch response components. It was chosen because both MEG- and EEG-derived pitch responses are prominent at frontocentral sites (e.g., Bidelman & Grall, 2014; Krishnan, Gandour, & Suresh, 2015b; Krumbholz, Patterson, Seither-Preisler, Lammertmann, & Lutkenhoner, 2003). It also allows us to compare our CPR data with Fz-derived POR data (e.g., Bidelman & Grall, 2014; Gutschalk, Patterson, Scherg, Uppenkamp, & Rupp, 2004). In addition, this electrode configuration was exploited to improve the signal-to-noise ratio of the CPR components by differentially amplifying (i) the non-inverted components recorded at Fz-linked(T7/T8) and (ii) the inverted components recorded at the temporal electrode sites (T7 and T8). For both averaging procedures, the analysis epoch was 1600 ms including the 100 ms pre-stimulus baseline.

#### 2.5. CPR magnitude

The evoked response to the entire three segment (noise-pitch-noise) stimulus is characterized by obligatory components (P1/N1) corresponding to the onset of energy in the precursor noise segment of the stimulus followed by three transient CPR components (Na, Pb, Nb) occurring after the onset of the pitch-eliciting segment of the stimulus, and an offset component (Po) that occurs after the offset of the last noise segment. We evaluated only the magnitude of the CPR components. Peak-to-peak amplitude of Na-Pb and Pb-Nb were measured to characterize the effects of changes in location of the turning point throughout the duration of the pitch contour. Peak-to-peak amplitude of Na-Pb and Pb-Nb were also measured at the temporal (T7/T8) electrode sites to evaluate response asymmetry. To enhance visualization of the asymmetry effects along a spectrotemporal dimension, a joint



time frequency analysis using a continuous wavelet transform was performed on the grand average waveforms derived from the temporal electrodes.

## 2.6. Statistical analysis

Separate, two-way (group x stimulus), mixed model ANOVAs (SAS®; SAS Institute, Inc., Cary, NC, USA) were performed on each component of peak-to-peak amplitude (Na-Pb, Pb-Nb) derived from the Fz electrode site; a three-way (group x stimulus x electrode site), mixed model ANOVA on peak-to-peak amplitude derived from the temporal electrode sites (T7/T8). Group (Chinese, English) functioned as the between-subjects factor; subjects nested within group served as a random factor. Stimulus (TP1–3) and electrode site (T7 [left]/T8 [right]) were treated as within-subject factors. All multiple pairwise comparisons were corrected with a Bonferroni significance level set at  $\alpha = 0.05$ . Partial eta-squared ( $\eta_p^2$ ) values, where appropriate, were reported to indicate effect sizes.

## 3. Results

### 3.1. Response morphology

Fig. 2 (top) illustrates that Fz-linked(T7/T8)-derived CPR components (Na, Pb, Nb) of the pitch-eliciting segment (color), preceded and followed by noise segments (black), are clearly identifiable embedded within the stimulus paradigm. No remarkable differences between groups are observed in their CPR responses during the noise segments (top: black segment). Na, Pb, and Nb (top: color segment) are the most robust pitch-relevant components. Fig. 2 (bottom) displays only the time window of the grand averaged CPR components for each turning point stimulus (TP1–3). CPR waveforms elicited by the three stimuli (bottom) show that the amplitude of the pitch-relevant components (Na, Pb, Nb) appear to be more robust for the Chinese as compared to the English, especially in response to stimuli with later location of the turning point (TP2–3).

### 3.2. Fz: amplitude (Na-Pb, Pb-Nb)

Fig. 3 displays mean peak-to-peak amplitude of CPR components (Na-Pb, Pb-Nb) for both language groups across the TP continuum. In the case of Na-Pb (top), ANOVA revealed main effects of group ( $F(1, 22) = 13.92, p = 0.0012, \eta_p^2 = 0.388$ ) and stimulus ( $F(2, 44) = 3.64, p = 0.0343, \eta_p^2 = 0.142$ ), and a group x stimulus interaction ( $F(2, 44) = 4.14, p = 0.0226, \eta_p^2 = 0.158$ ). By group, post hoc multiple comparisons indicated that amplitude of the Chinese group in response to TP2–3 was larger than TP1 (TP1 vs. TP2:  $t = -3.35, p = 0.0049$ ; TP1 vs. TP3:  $t = -3.44, p = 0.0038$ ). Later temporal location of the turning point simultaneously results in decreasing acceleration rates in the falling section of the pitch contour and increasing acceleration rates in the rising section. In contrast, amplitude of the English group was comparatively invariant across the continuum. None of their paired comparisons reached significance. By stimulus, Chinese listeners, relative to English, exhibited larger amplitude evoked by TP2–3 (TP2:  $t = 4.00, p < 0.0002$ ; TP3:  $t = 3.83, p < 0.0004$ ).

In the case of Pb-Nb (bottom), ANOVA similarly revealed main effects of group ( $F(1, 22) = 6.23, p = 0.0205, \eta_p^2 = 0.221$ ) and stimulus ( $F(2, 44) = 19.92, p < 0.0001, \eta_p^2 = 0.475$ ), and a group x stimulus interaction ( $F(2, 44) = 3.93, p = 0.0269, \eta_p^2 = 0.152$ ). By group, post hoc multiple comparisons indicated that amplitude of the Chinese group in response to TP2–3 was larger than TP1 (TP2,  $t = -5.92, p < 0.0001$ ; TP3,  $t = -4.52, p = 0.0001$ ). TP2 was also larger in amplitude than TP1 in the English group ( $t = -2.95, p = 0.0151$ ). By stimulus, amplitude was larger in Chinese than English listeners in response to TP2 ( $t = 2.25, p = 0.0297$ ) and TP3 ( $t = 3.01, p = 0.0044$ ).

### 3.3. T7/T8: amplitude of CPR components

Grand average waveforms of the CPR components (left columns) in response to each stimulus along the turning point continuum per language group and their corresponding spectra (right columns) are displayed in Fig. 4. A rightward asymmetry (T8 > T7) is apparent in the Chinese group as evidenced by CPR components in both waveforms and spectrotemporal plots more or less across the continuum—especially evident in TP2 and TP3. The English group, on the other hand, shows the opposite tendency toward a leftward asymmetry (T7 > T8). These language dependent effects are clearly evident in the spectrotemporal plots of TP2 and TP3.

T7/T8 peak-to-peak amplitude of Na-Pb (top) and Pb-Nb (bottom) per turning point stimulus are displayed for each language group and electrode site in Fig. 5. As reflected by Na-Pb, the three-way (group x stimulus x electrode site) ANOVA revealed a main effect of stimulus ( $F(2, 44) = 5.27, p < 0.0089, \eta_p^2 = 0.193$ ), electrode site ( $F(1, 66) = 4.20, p < 0.0443, \eta_p^2 = 0.060$ ), and a group x electrode site interaction ( $F(1, 66) = 7.68, p < 0.0073, \eta_p^2 = 0.104$ ). Pooling across stimulus, post hoc comparisons showed a leftward asymmetry (T7 > T8) in the English group ( $t_{66} = 3.41, p = 0.0011$ ). Chinese listeners' response amplitude was larger than English over the right temporal site only ( $t_{66} = 2.65, p = 0.0100$ ). Pooling across group and temporal electrode site, the stimulus main effect showed that TP2 amplitude was larger than TP1 ( $t_{44} = -3.24, p = 0.0068$ ).

As reflected by Pb-Nb, the three-way (group x stimulus x electrode site) ANOVA revealed a main effect of stimulus ( $F(2, 44) = 19.40, p < 0.0001, \eta_p^2 = 0.469$ ), a group x electrode site interaction ( $F(1, 66) = 5.63, p = 0.0206, \eta_p^2 = 0.079$ ), and a stimulus x electrode site interaction ( $F(2, 66) = 3.52, p = 0.0351, \eta_p^2 = 0.097$ ). Pooling across stimulus, the group x electrode site interaction showed that amplitude of the Chinese group was larger over the right electrode site relative to the left ( $t_{66} = -2.36, p = 0.0211$ ). Pooling across group, the stimulus x electrode site interaction showed that TP2 amplitude was larger over the right as compared to the left temporal site ( $t_{66} = -2.27, p < 0.0266$ ). Over the left temporal site, TP2 amplitude was larger than TP1 ( $t_{66} = -3.89, p = 0.0007$ ); over the right temporal site, amplitude of TP2 and TP3 were both larger than TP1 (TP1 vs. TP2:  $t_{66} = -6.68, p < 0.0001$ ; TP1 vs. TP3:  $t_{66} = -3.41, p = 0.0033$ ) and, in addition, TP2 amplitude was larger than TP3 ( $t_{66} = 3.26, p < 0.0052$ ).



Thus, the effects of language experience on sensitivity to location of the turning point implicate the right temporal site. Across language groups, TP2–3 amplitude is larger than TP1 over the right temporal site; TP2 amplitude is larger than TP3. Over the left temporal site, TP2 is larger than TP1.

## 4. Discussion

The major findings of this study demonstrate that pitch-related neural activity—as reflected in the amplitude of the CPR response components (Na-Pb, and Pb-Nb)—exhibit differential sensitivity to changes in temporal location of the turning point for English and Chinese listeners. In the Chinese group, Na–Pb and Pb–Nb amplitude were selectively enhanced at the Fz electrode site for pitch contours with later turning points (TP2, temporal location of turning point at 46% of total duration); TP3, 62%) relative to one with an early turning point (TP1, 30%). Language-dependent enhancement ( $C > E$ ) is restricted to pitch contours TP2 and TP3 for Na–Pb and Pb–Nb. Over the right temporal electrode site (T8), Na–Pb amplitude is larger in the Chinese group relative to the English. In contrast, a leftward asymmetry ( $T7 > T8$ ) is found in the English group. As reflected by Pb–Nb, a rightward asymmetry ( $T8 > T7$ ) is observed in the Chinese group. These findings suggest that the neural mechanism(s) underlying processing of pitch in the right auditory cortex reflect experience-dependent modulation of sensitivity to differentially weighted integration of changes in acceleration rates of rising and falling sections and the temporal location of the turning point.

### 4.1. Cortical pitch response components are sensitive to differential weighting of acceleration and temporal location of turning point

The major finding of this study is that language experience enhances cortical pitch response components as a function of differential weighting of the pitch-related neural activity to changes in acceleration rates of the falling and rising sections and the temporal location of the turning point within bidirectional pitch contours. In such a scheme, an optimal response is predicted when the weights of all three acoustic attributes (rising and falling acceleration rates; temporal location of turning point) closely approximate those of native exemplars. Among the three TP bidirectional pitch contours, an optimally weighted response would be expected to occur when they closely approximate the temporal attributes of Mandarin Tone 2, a high falling-rising pitch contour. Consistent with this notion, language-dependent enhancement of pitch-related neural activity is observed for both CPR Fz components (Na–Pb, Pb–Nb) in response to TP2 and TP3, the two stimuli with temporal attributes that most closely approximate Tone 2. For both groups, smaller amplitudes evoked by TP1 likely reflect a language-universal effect triggered by low weightings on all three components. In contrast to TP2 and TP3, TP1 is distinguished by the relatively early temporal location of the turning point (74 ms) in conjunction with a steep, falling slope from pitch onset to turning point ( $F_0$ , 40 Hz; velocity,  $-69.1$  st/s). Results of perception tests, however, indicate that TP1 is easily identified as Tone 2 (Moore & Jongman, 1997). But neural responses at this early level of early sensory processing may not be extrapolated directly to a speech perception task that involves attention, working memory, and a categorical, decision-making judgment. In this preattentive study, participants were free of task requirements other than to watch a muted video during data acquisition. Moreover, the time scales of processing are not

comparable. After presentation of a stimulus in the speech perception task, there was an inter-trial interval of 2 s to allow time for making a decision about its tonal identity. In contrast, CPR responses are elicited automatically within 0.5 s at an early cortical sensory stage of processing. As demonstrated by the CPR, the upper and/or lower bounds of sensitivity to temporal attributes of pitch contours are narrower than those associated with a behavioral identification task. In this experiment, the location of the turning point was varied in the context of a fixed pitch onset and offset. Further research is warranted to more fully illuminate the differential sensitivity of CPR components (i) by varying pitch onset in the context of a fixed turning point and offset and (ii) by varying pitch offset in the context of a fixed turning point and onset.

Indeed, as indexed by Pb–Nb over the right temporal site, we observe that TP2 and TP3 are larger in amplitude than TP1 across language groups. In addition, TP2 exhibits larger amplitude than TP3. The latter is characterized by the relatively later temporal location of the turning point (156 ms) in conjunction with a steep, rising slope from turning point to pitch offset ( $F_0$ , 50 Hz; velocity, 94.8 st/s). TP2, on the other hand, exhibits relatively shallow slopes in both sections (falling, rising) of the pitch contour. The absence of a language experience effect over the right temporal site, as reflected by Pb–Nb, emphasizes the sensory nature of the CPR.

At the level of the brainstem, both Chinese and English listeners exhibit a systematic reduction in response amplitude with increasing acceleration rates (Krishnan, et al., 2010). Pitch representations of Chinese listeners, however, are more resistant to degradation. This experience-dependent advantage extends even to acceleration rates that fall proximal to or outside the boundary of natural speech. At the cortical level, CPR components (Na–Pb, Pb–Nb) reveal that Mandarin natives show a decrease in amplitude with increasing pitch acceleration across three variants of Tone 2 within the normal voice pitch range (Krishnan, et al., 2014). In a subsequent crosslanguage study (Chinese vs. English; Krishnan, Gandour, et al., 2015a), the stimulus set consisted of four variants of Tone 2 representing a range of acceleration rates: two within the normal voice pitch range; two outside. Chinese listeners show an experience-dependent response enhancement restricted to just those pitch contours with acceleration rates that fall within the bounds of one's native language. This reduced sensitivity to rapid pitch acceleration presumably reflects adaptation, and recovery from adaptation-related degradation of neural activity associated with rapidly changing pitch sections (e.g., neural desynchronization, synaptic inefficiency).

CPR components, Na–Pb and Pb–Nb, are sensitive to different temporal attributes of pitch. For example, the Na–Pb time window optimally represents neural processing relevant to pitch salience (Krishnan, Gandour, & Suresh, 2016). But salience was fixed throughout the duration of all pitch contours. It is therefore not surprising that experience-dependent modulation of pitch salience is restricted to this earlier time window. As reflected by Na–Pb and Pb–Nb at the Fz site, the experience-dependent effect is restricted to TP2 and TP3. In agreement with recent CPR studies (Krishnan, Gandour, Ananthakrishnan, et al., 2015; Krishnan, Gandour, et al., 2015a, 2015b), this may indicate increased selectivity to a dynamic temporal attribute, pitch acceleration, in both temporal integration windows. This explanation implies that experience-dependent effects are targeted to specific temporal

integration windows in which optimal processing occurs for a particular dimension of pitch. Thus, pitch processing involves a hierarchy of both sensory and extrasensory effects whose relative weighting varies depending on both language experience and the sensitivity of neural activity within a given temporal integration window to particular attributes of pitch.

More broadly, we infer that extrasensory processes are overlaid on sensory processes to modulate long-term, experience-driven, adaptive pitch mechanisms at early sensory levels of pitch processing in the auditory cortex as well as in the brainstem by enhancing response selectivity of neural elements to behaviorally-relevant temporal attributes of native pitch contours. By extrasensory, we mean neural processes at a higher hierarchical level beyond the purely sensory processing of acoustic attributes of the stimulus. One likely candidate for stored representations of pitch attributes at this early sensory cortical level of processing is analyzed sensory memory (Cowan, 1988; Xu, Gandour, & Francis, 2006). In contrast to traditional encapsulated stored memory, Hasson et al. (2015) propose a biologically-motivated process memory framework in which cortical neural circuits integrate past information with incoming information. Process memory refers to the integration of active traces of past information that are used by a neural circuit to process incoming information in the present moment. By the Hasson et al. model, our CPR responses would be activated in the early stages of this processing memory hierarchy and utilize short temporal receptive windows where the neural dynamics are more rapid.

#### **4.2. Hemispheric preferences in pitch processing differ between Chinese and English listeners**

Our findings reveal differences in relative asymmetry over the temporal electrode sites depending upon language experience. As indexed by Na–Pb and Pb–Nb, it is only over the right temporal site (T8) that the amplitude of these components are larger for Chinese listeners as compared to English. These findings converge with an extant literature that highlights the more fine-grained resolution in the right hemisphere for processing linguistic as well as nonlinguistic attributes of pitch (Friederici, 2011; Friederici & Alter, 2004; Hyde, Peretz, & Zatorre, 2008; Jamison, Watkins, Bishop, & Matthews, 2006; Johnsrude, Penhune, & Zatorre, 2000; Meyer, 2008; Schonwiesner, Rubsamen, & von Cramon, 2005; Zatorre, Belin, & Penhune, 2002; Zatorre & Gandour, 2008). Important to bear in mind, however, is that task-dependent functional imaging studies aggregate across multiple time windows in the pitch processing hierarchy. Later time windows recruit regions not restricted to primary auditory cortex. Indeed, stronger asymmetries develop as pitch-related neural activity moves out of the primary auditory cortex for more complex pitch stimuli (Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002).

A simple invocation of more fine-grained processing for pitch in the right hemisphere would predict a relative rightward asymmetry regardless of language experience. Yet it is only the Chinese group that exhibits a rightward asymmetry. Our CPR components reflect early sensory, task-free, pitch-specific neural activity in the primary auditory cortex. Using the CPR, we have found that language-dependent effects apply only to pitch contours with temporal attributes that occur in natural speech. We infer that such effects result from experience-driven, extrasensory modulation of adaptive pitch mechanisms at early sensory

levels of pitch processing in the right auditory cortex. In this case, these adaptive mechanisms afford more fine-grained processing of linguistically-relevant, temporal pitch attributes. Such a formulation is consistent with our recent observations of rightward asymmetry at temporal electrode sites (Krishnan, et al., 2014, time-variant vs. time-invariant; pitch direction & location of peak acceleration, Krishnan, Gandour, et al., 2015a; Krishnan, Gandour, et al., 2015b, rate of pitch acceleration; Krishnan, et al., 2016, salience).

Electrophysiological studies of categorical perception of Mandarin lexical tones (Tone 2, Tone 4) have revealed that across-category deviants, relative to within-category deviants, elicited larger oddball responses in the left recording sites than in the right sites (MMN: Xi, Zhang, Shu, Zhang, & Li, 2010; N2b, P3b: Zhang, Xi, Wu, Shu, & Li, 2012). Using the same stimuli with functional magnetic resonance imaging, Zhang, et al. (2011) conducted region-of-interest analyses in the left middle temporal gyrus (MTG) and right lateral Heschl's gyrus (HG). Across-category deviants elicited stronger activation than within-category deviants in the left MTG; just the opposite was observed in the right HG. These previous findings are consistent with our claim that the observed, rightward asymmetry over the temporal sites—as elicited by TP1, TP2, and TP3—does not represent a change in tonal category between Tone 2 and Tone 3.

Finally, we observed a relative leftward asymmetry for the English group characterized by a reduction in the response magnitude—an effect opposite to the Chinese group—over the right temporal electrode site compared to the left electrode site. None of our previous experiments in which we manipulated pitch acceleration revealed any asymmetries for the English group. We are at a loss to advance any plausible explanation. We can only speculate that pitch mechanisms in the right hemisphere for English monolinguals may be more susceptible to degradation resulting from manipulation of both turning point and pitch acceleration.

#### 4.3. Predictive coding may underlie experience-dependent processing of pitch

Growing evidence shows pitch-related neural activity in both the primary auditory cortex and medial Heschl's gyrus (HG) as well as in the adjacent more lateral non-primary areas of HG, suggesting that pitch-relevant information is available in multiple areas of the auditory cortex. A hierarchical processing framework for coordinated interaction between these areas is provided by application of a predictive coding model of pitch perception to depth-electrode recordings of pitch-relevant neural activity along HG (Kumar & Schonwiesner, 2012; Kumar, et al., 2011; see Rao & Ballard, 1999, for details of the model). Essentially, Kumar, et al. (2011) have shown that the pitch prediction at the higher level is continually adjusted in a recursive manner by changing the strengths of the top down and bottom up connections to optimize pitch representation.

Since the EEG derived CPR data presented here is essentially a far field version of the depth recorded pitch onset responses utilized by Kumar, et al. (2011), it is reasonable to capture our results within this framework. Applied to our data, this framework suggests that CPR changes attributable wholly to acoustic properties of the stimulus invoke a recursive process in the representation of pitch (initial pitch prediction, error generation, error correction). At this level, the hierarchical flow of processing and its connectivity strengths along the HG are

essentially the same regardless of one's language background. However, the initial pitch prediction at the level of the lateral HG is more precise for Chinese because of their access to stored information about pitch contours—their temporal attributes to be more precise—that represent best approximations to native pitch contours (TP2, TP3) with a smaller error term. Consequently, the top-down connection from lateral HG to medial and middle HG is stronger than the bottom-up connection. The opposite would be true for English because of their less precise initial prediction. Language experience therefore alters the nature of the interaction between levels along the hierarchy of pitch processing by modulating connection strengths. The hierarchical processing memory framework is broadly consistent with predictive coding.

Pitch processing in the auditory cortex is influenced by inputs from subcortical structures that are themselves subject to experience-dependent plasticity. It is likely that top-down connections in the hierarchy provide feedback to adjust the effective time scales of processing at each stage to optimally control the temporal dynamics of pitch processing. Language-dependent changes in the CPR by Chinese may reflect interplay between sensory and extrasensory processing. This expanded model represents a unified, physiologically plausible, theoretical framework that includes both cortical and subcortical components in the hierarchical processing of pitch.

#### 4.4. Conclusions

The language-dependent enhancement of pitch relevant neural activity to bidirectional pitch contours approximating temporal attributes of native pitch contours may reflect optimal, differentially weighted, integration of pitch relevant neural activity to changes in acceleration rates of the falling and rising sections and the temporal location of the turning point. The relatively greater stimulus selectivity of Pb-Nb compared to Na-Pb appears to indicate that experience-dependent effects are targeted to specific temporal integration windows in which optimal processing occurs for a particular attribute of pitch. Thus, pitch processing involves a hierarchy of both sensory and extrasensory effects whose relative weighting varies depending on both language experience and the sensitivity of neural activity within a given temporal integration window to particular attributes of pitch. Language-dependent differences in hemispheric asymmetry at temporal electrode sites may implicate fundamental differences in early cortical processing of pitch-relevant information in dynamic pitch contours.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

This work was supported by the National Institutes of Health 5R01DC008549-7 (A.K.). Thanks to Rongrong Zhang for her assistance with statistical analysis (Department of Statistics); Breanne Lawler, Whitney Lyle, and Kari Russell for their help with data acquisition. The authors declare no conflict of interest.

This work was supported by National Institutes of Health Grant No. 5R01DC008549-7

## References

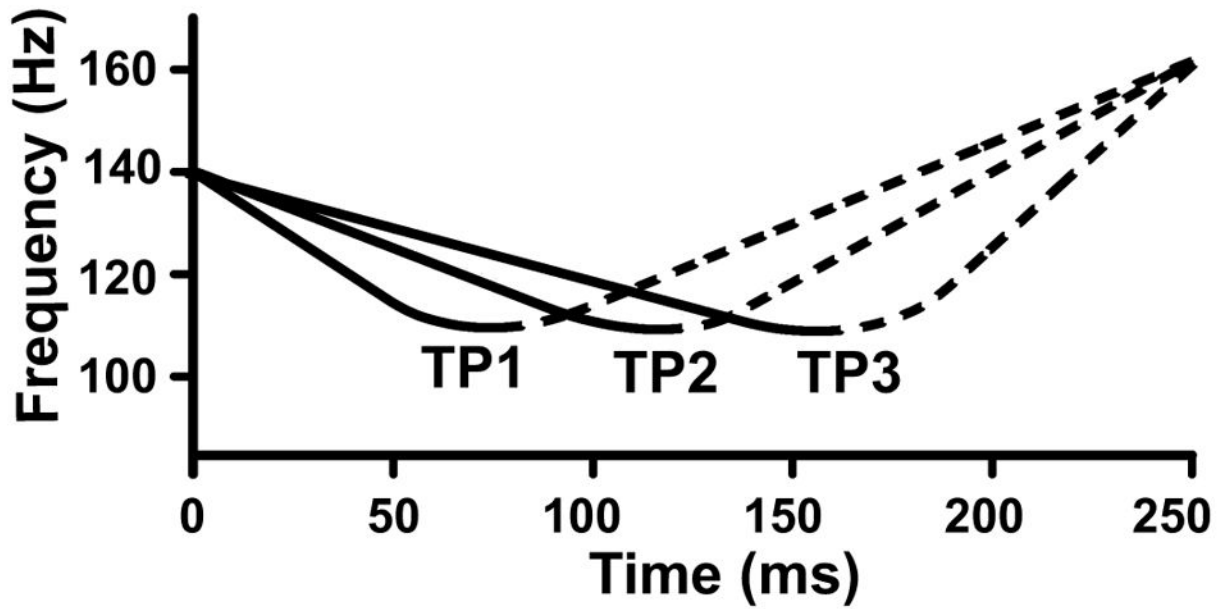
- Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*. 2011; 23:425–434. [PubMed: 19925180]
- Bidelman GM, Grall J. Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *Neuroimage*. 2014; 101:204–214. [PubMed: 25019679]
- Bidelman GM, Lee CC. Effects of language experience and stimulus context on the neural organization and categorical perception of speech. *Neuroimage*. 2015; 120:191–200. [PubMed: 26146197]
- Cowan N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*. 1988; 104:163–191. [PubMed: 3054993]
- Friederici AD. The brain basis of language processing: from structure to function. *Physiological Reviews*. 2011; 91:1357–1392. [PubMed: 22013214]
- Friederici AD, Alter K. Lateralization of auditory language functions: a dynamic dual pathway model. *Brain and Language*. 2004; 89:267–276. [PubMed: 15068909]
- Gandour, JT., Krishnan, A. Processing tone languages. In: Hickok, G., Small, SL., editors. *Neurobiology of language*. New York: Academic Press; 2016. p. 1095-1107.
- Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A. Temporal dynamics of pitch in human auditory cortex. *Neuroimage*. 2004; 22:755–766. [PubMed: 15193604]
- Hasson U, Chen J, Honey CJ. Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*. 2015; 19:304–313. [PubMed: 25980649]
- Hyde KL, Peretz I, Zatorre RJ. Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*. 2008; 46:632–639. [PubMed: 17959204]
- Jamison HL, Watkins KE, Bishop DV, Matthews PM. Hemispheric specialization for processing auditory nonspeech stimuli. *Cerebral Cortex*. 2006; 16:1266–1275. [PubMed: 16280465]
- Jia S, Tsang YK, Huang J, Chen HC. Processing Cantonese lexical tones: Evidence from oddball paradigms. *Neuroscience*. 2015; 305:351–360. [PubMed: 26265553]
- Johnson KL, Nicol TG, Kraus N. Brain stem response to speech: A biological marker of auditory processing. *Ear and Hearing*. 2005; 26:424–434. [PubMed: 16230893]
- Johnsrude IS, Penhune VB, Zatorre RJ. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*. 2000; 123:155–163. [PubMed: 10611129]
- Kraus N, Banai K. Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*. 2007; 16:105–110.
- Krishnan A, Gandour JT. The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*. 2009; 110:135–148. [PubMed: 19366639]
- Krishnan A, Gandour JT. Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: electrophysiological evidence. *Acoustics Australia*. 2014; 42:166–178. [PubMed: 25838636]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V. Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*. 2014; 59:1–12. [PubMed: 24751993]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V. Language experience enhances early cortical pitch-dependent responses. *Journal of Neurolinguistics*. 2015; 33:128–148. [PubMed: 25506127]
- 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 and Language*. 2010; 114:193–198. [PubMed: 20570340]
- Krishnan A, Gandour JT, Suresh CH. Experience-dependent enhancement of pitch-specific responses in the auditory cortex is limited to acceleration rates in normal voice range. *Neuroscience*. 2015a; 303:433–445. [PubMed: 26166727]



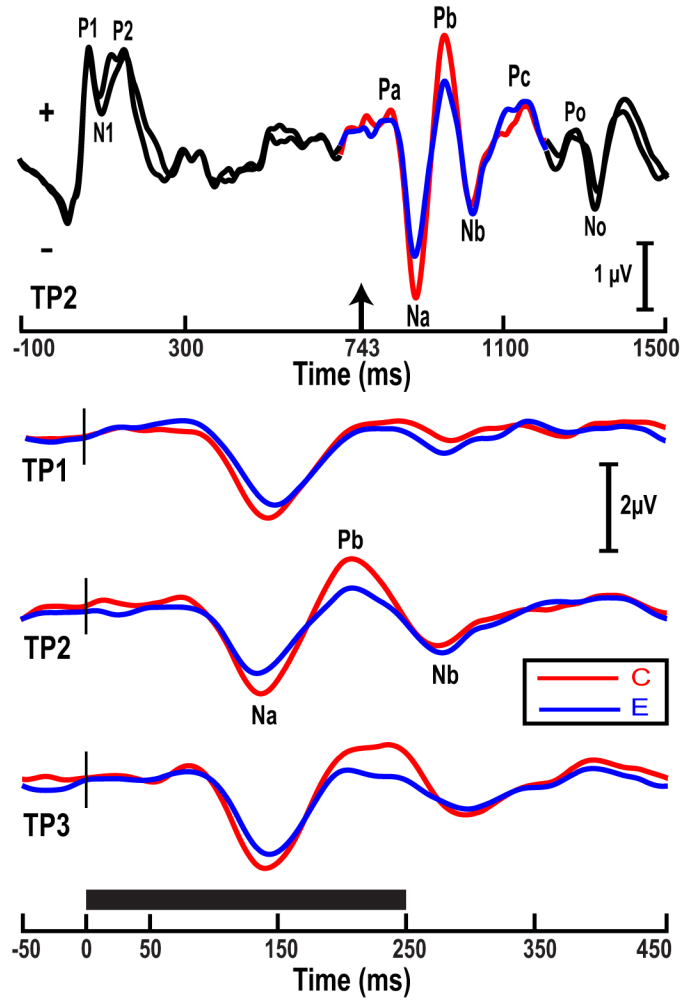
- Krishnan A, Gandour JT, Suresh CH. Pitch processing of dynamic lexical tones in the auditory cortex is influenced by sensory and extrasensory processes. *European Journal of Neuroscience*. 2015b; 41:1496–1504. [PubMed: 25943576]
- Krishnan A, Gandour JT, Suresh CH. Language-experience plasticity in neural representation of changes in pitch salience. *Brain Research*. 2016; 1637:102–117. [PubMed: 26903418]
- Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*. 2009; 21:1092–1105. [PubMed: 18702588]
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lutkenhoner B. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*. 2003; 13:765–772. [PubMed: 12816892]
- Kumar S, Schonwiesner M. Mapping human pitch representation in a distributed system using depth-electrode recordings and modeling. *Journal of Neuroscience*. 2012; 32:13348–13351. [PubMed: 23015425]
- Kumar S, Sedley W, Nourski KV, Kawasaki H, Oya H, Patterson RD, Howard MA 3rd, Friston KJ, Griffiths TD. Predictive coding and pitch processing in the auditory cortex. *Journal of Cognitive Neuroscience*. 2011; 23:3084–3094. [PubMed: 21452943]
- Laver, J. Principles of phonetics. New York, NY: Cambridge University Press; 1994. (Chapter)
- Li P, Sepanski S, Zhao X. Language history questionnaire: A web-based interface for bilingual research. *Behavioral Research Methods*. 2006; 38:202–210.
- Meyer M. Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitschrift für Neuropsychologie*. 2008; 19:101–115.
- Moore CB, Jongman A. Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*. 1997; 102:1864–1877. [PubMed: 9301064]
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Patel AD, Iversen JR. The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*. 2007; 11:369–372. [PubMed: 17698406]
- Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. *Neuron*. 2002; 36:767–776. [PubMed: 12441063]
- Prom-on S, Xu Y, Thipakorn B. Modeling tone and intonation in Mandarin and English as a process of target approximation. *Journal of the Acoustical Society of America*. 2009; 125:405–424. [PubMed: 19173427]
- Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*. 1999; 2:79–87. [PubMed: 10195184]
- Schonwiesner M, Rubsamen R, von Cramon DY. Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *European Journal of Neuroscience*. 2005; 22:1521–1528. [PubMed: 16190905]
- Somers M, Aukes MF, Ophoff RA, Boks MP, Flier W, de Visser KC, Kahn RS, Sommer IE. On the relationship between degree of hand-preference and degree of language lateralization. *Brain and Language*. 2015; 144:10–15. [PubMed: 25880901]
- Swaminathan J, Krishnan A, Gandour JT. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*. 2008; 19:1163–1167. [PubMed: 18596621]
- Tsang YK, Jia S, Huang J, Chen HC. ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neuroscience Letters*. 2011; 487:268–272. [PubMed: 20970477]
- Tzounopoulos T, Kraus N. Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron*. 2009; 62:463–469. [PubMed: 19477149]
- Wong PC, Perrachione TK. Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*. 2007; 28:565–585.
- Xi J, Zhang L, Shu H, Zhang Y, Li P. Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*. 2010; 170:223–231. [PubMed: 20633613]
- Xu Y. Contextual tonal variations in Mandarin. *Journal of Phonetics*. 1997; 25:61–83.

- Xu Y, Gandour JT, Francis AL. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*. 2006; 120:1063–1074. [PubMed: 16938992]
- Xu, Y., Prom-on, S. PENTAtainer1.praat. 2010–2015. Available from <http://www.homepages.ucl.ac.uk/~uclyyix/PENTAtainer1/>
- Xu Y, Sun X. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*. 2002; 111:1399–1413. [PubMed: 11931317]
- Yost WA. Pitch of iterated rippled noise. *Journal of the Acoustical Society of America*. 1996a; 100:511–518. [PubMed: 8675844]
- Yost WA. Pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America*. 1996b; 100:3329–3335. [PubMed: 8914314]
- Zatorre RJ, Baum SR. Musical melody and speech intonation: Singing a different tune. *PLoS Biology*. 2012; 10:e1001372. [PubMed: 22859909]
- Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*. 2002; 6:37–46. [PubMed: 11849614]
- Zatorre RJ, Gandour JT. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*. 2008; 363:1087–1104. [PubMed: 17890188]
- Zhang L, Xi J, Wu H, Shu H, Li P. Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *Neuroreport*. 2012; 23:35–39. [PubMed: 22113214]
- Zhang L, Xi J, Xu G, Shu H, Wang X, Li P. Cortical dynamics of acoustic and phonological processing in speech perception. *PloS One*. 2011; 6:e20963. 2011/06/23 ed. [PubMed: 21695133]

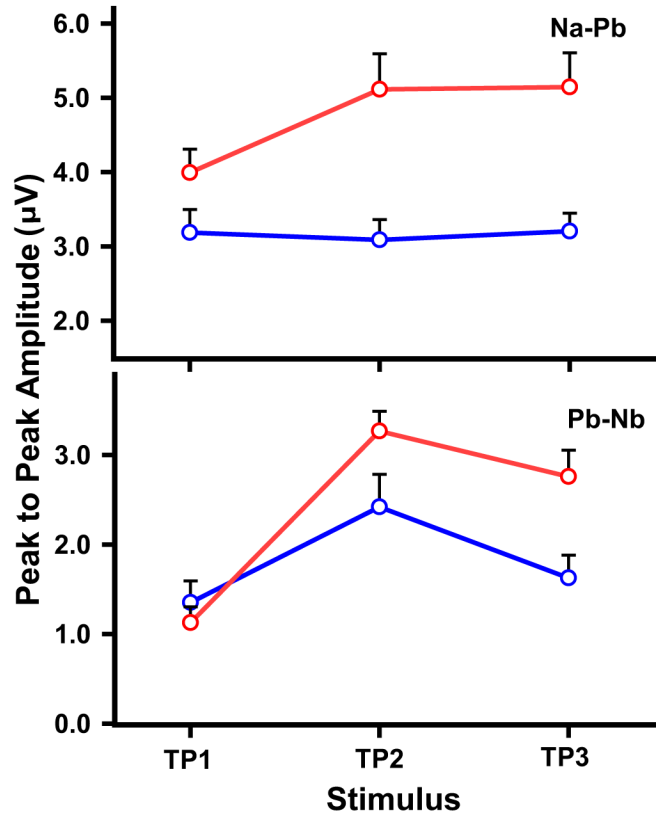
- Cortical pitch-specific responses exhibit sensitivity to location of turning point
- Crosslanguage effects are restricted to pitch contours with later turning points
- Crosslanguage effects are observed at both frontal and temporal electrode sites
- Language-dependent sensitivity to turning point is limited to right temporal site
- Experience modulates sensitivity to falling/rising acceleration and turning point



**Figure 1.** IRN stimuli used to acquire CPR responses to dynamic, curvilinear F0 contours varying in location of the turning point. Each stimulus consists of two segments differing in direction: falling (solid), rising (dashed).



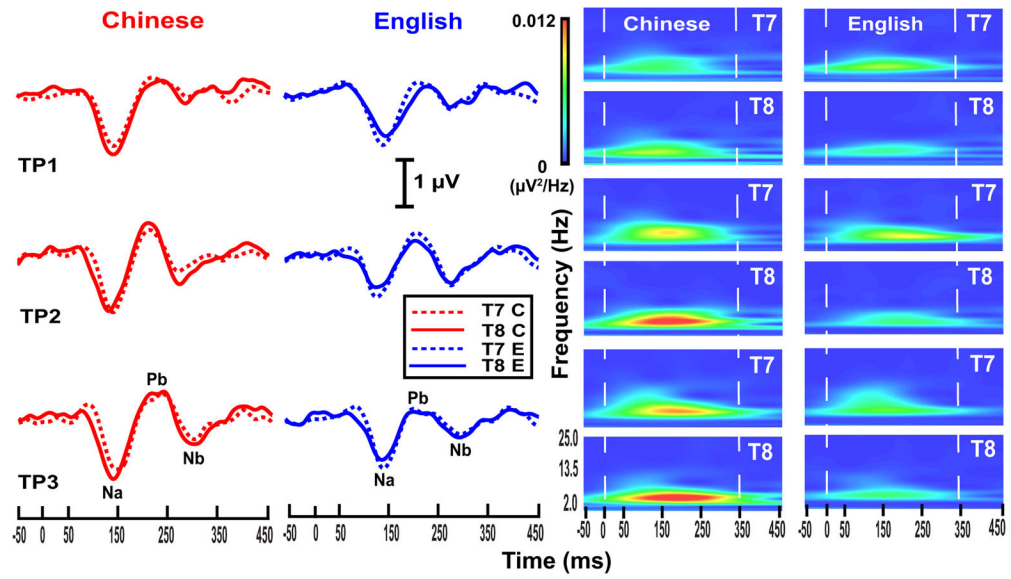
**Figure 2.** Grand average waveforms of the Chinese (C) and English (E) groups at the Fz electrode site per pitch stimulus (TP1–3). Waveforms consisting of three segments (top) illustrate the experimental paradigm used to acquire cortical responses: a 250 ms pitch segment (IRN,  $n = 32$ ) preceded and followed by 750 ms and 250 ms noise segments, respectively. Solid black horizontal bar indicates the duration of each stimulus. The up arrow at 743 ms (top) marks the onset of the pitch-eliciting segment of the stimulus; short vertical stroke (black) marks the time point in the waveforms. See Section 3.1 for details.



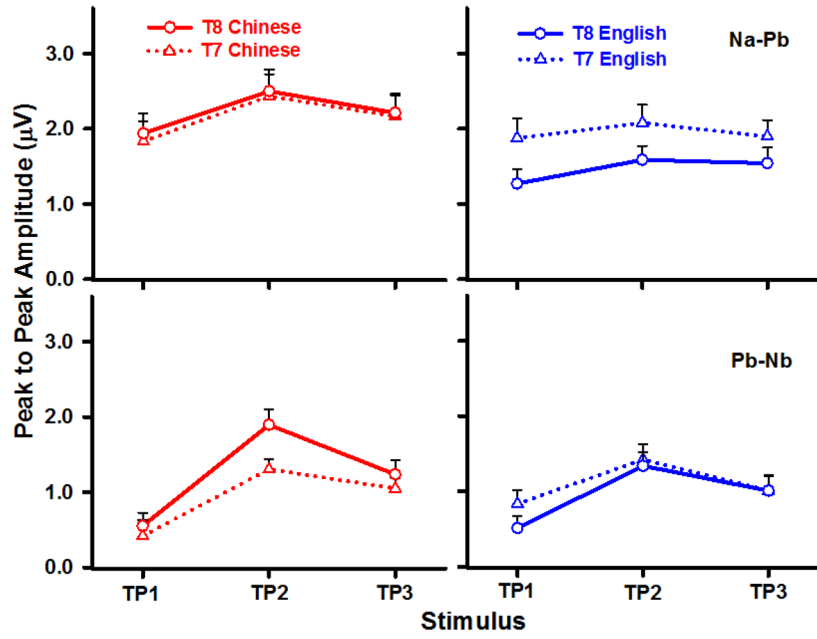
**Figure 3.**

Mean peak-to-peak amplitude of CPR components (Na-Pb, Pb-Nb) elicited by each of the three TP stimuli at the Fz electrode site in both Chinese and English groups. For Na-Pb (top), TP2 and TP3 are larger in amplitude than TP1 for the Chinese group only. Amplitude is larger for the Chinese group relative to the English for TP2–3. Similarly for Pb-Nb (bottom), the Chinese group advantage is evident for TP2–3. See Section 3.2 for details.





**Figure 4.** Grand average waveforms (left) and their corresponding spectra (right) of the CPR components for the two language groups (Chinese, red; English, blue) recorded at electrode sites T7 (dashed) and T8 (solid) for each of the three stimuli (TP1–3). Both Na-Pb and Pb-Nb time windows are included between two, vertical white dashed lines (right). The zero on the x-axis denotes the time of onset of the pitch-eliciting segment. See Section 3.3 for details.



**Figure 5.** Mean peak-to-peak amplitude of CPR components (Na-Pb, Pb-Nb) elicited by each of the three TP stimuli at the T7/T8 temporal electrode sites in both Chinese and English groups. In the case of Na-Pb (top row), amplitude is larger in the Chinese group relative to the English in the right temporal site only. A leftward asymmetry (T7 > T8), on the other hand, is observed for the English group. In the case of Pb-Nb, the Chinese group exhibits a rightward asymmetry (T8 > T7) for TP2 only. See Section 3.3 for details.

**Table 1**

Duration values per section of TP pitch contours

Stimulus	F0 onset to turning point	Turning point to F0 offset
TP1	48	148.48
TP2	87.2	112.53
TP3	137.5	60.03

Note. Duration values are expressed in ms.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

TP stimuli compared to mean speed of pitch change (Xu & Sun, 2002)

Stimulus	Mean velocity (st/s)	
	Fall -55@4 st interval	Rise 86@7 st interval
TP1	-69.1	43.1
TP2	-41.2	59.1
TP3	-29.5	94.8

Note. Gray-highlighted cells indicate that mean velocity (st/s) of falling or rising section of the TP stimulus does not exceed the mean fall or rise velocity reported for two pitch intervals (4 and 7 st) (Xu & Sun, 2002, Table VII).