

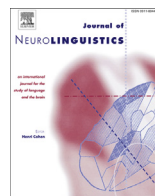


ELSEVIER

Contents lists available at ScienceDirect

Journal of Neurolinguistics

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



CrossMark

Language experience enhances early cortical pitch-dependent responses

Ananthanarayan Krishnan^a, Jackson T. Gandour^{a,*},
Saradha Ananthakrishnan^{a,1},
Venkatakrishnan Vijayaraghavan^b

^a Department of Speech Language Hearing Sciences, Purdue University, USA

^b School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA

ARTICLE INFO

Article history:

Received 2 March 2014

Received in revised form 2 March 2014

Accepted 5 August 2014

Available online 1 September 2014

Keywords:

Pitch

Iterated rippled noise

Cortical pitch response

Tone language

Experience-dependent plasticity

Functional asymmetry

ABSTRACT

Pitch processing at cortical and subcortical stages of processing is shaped by language experience. We recently demonstrated that specific components of the cortical pitch response (CPR) index the more rapidly-changing portions of the high rising Tone 2 of Mandarin Chinese, in addition to marking pitch onset and sound offset. In this study, we examine how language experience (Mandarin vs. English) shapes the processing of different temporal attributes of pitch reflected in the CPR components using stimuli representative of within-category variants of Tone 2. 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 compared to English listeners for stimuli that fell within the range of Tone 2 citation forms. Discriminant function analysis revealed that the Na–Pb component was more than twice as important as Pb–Nb in grouping listeners by language affiliation. In addition, a stronger stimulus-dependent, rightward asymmetry was observed for the Chinese group at the temporal, but not frontal, electrode sites. This finding may reflect selective recruitment of experience-dependent, pitch-specific mechanisms in right auditory cortex to extract more complex, time-varying pitch patterns. Taken together,

* Corresponding author. Purdue University, Department of Speech Language Hearing Sciences, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038, USA. Tel.: +1 765 494 3821; fax: +1 765 494 0771.

E-mail addresses: rkkrish@purdue.edu (A. Krishnan), gandour@purdue.edu (J.T. Gandour), sananthakrishnan@towson.edu (S. Ananthakrishnan), vvijayar@purdue.edu (V. Vijayaraghavan).

¹ Saradha Ananthakrishnan is now affiliated with Towson University in the Department of Audiology, Speech-Language Pathology and Deaf Studies, Towson, MD, USA.

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 may vary depending on the relative linguistic importance of specific temporal attributes of dynamic pitch.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

1. Introduction

Pitch is a salient perceptual attribute that plays an important role in language and music (Oxenham, 2012; Plack, Oxenham, & Fay, 2005). Despite similarities in pitch processing between domains, empirical evidence supports the view that neural representations of pitch may be shaped by its functional properties in a given domain of expertise. Tone languages are especially useful for studying the effects of functional properties of pitch that are phonemic at the syllable level (Maddieson, 1978; Yip, 2002). It is well established that dynamic variations in voice fundamental frequency (F_0) provide the dominant acoustic cue for tonal recognition (Abramson, 1962; Gandour, 1994; Klatt, 1973; Xu, 2001). In the case of lexical tone, several cross-language (or cross-domain) studies have revealed experience-dependent neural plasticity at both cortical and subcortical levels of the brain (see Gandour, 2006; Gandour & Krishnan, 2014; Krishnan, Gandour, & Bidelman, 2012; Zatorre & Baum, 2012; Zatorre & Gandour, 2008, for reviews). 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.

Pitch is a multidimensional perceptual attribute that relies on several acoustic dimensions. In particular, F_0 height and contour (i.e., nonlinear change in pitch between onset and offset) have been revealed to be important, experience-dependent dimensions of pitch underlying the perception of lexical tone (Francis, Ciocca, Ma, & Fenn, 2008; Gandour, 1983; Gandour & Harshman, 1978; Huang & Johnson, 2011; Khouw & Ciocca, 2007). These same pitch dimensions have been targeted in recent studies of tonal processing in the human brain. Using the mismatch negativity (MMN), Chinese listeners, relative to English, were more sensitive to pitch contour than pitch height in response to Mandarin tones, indicating that MMN may serve as a neural index of the relative saliency of underlying dimensions of pitch that are differentially weighted by language experience (Chandrasekaran, Gandour, & Krishnan, 2007). In Cantonese, the magnitude and latency of MMN were sensitive to the size of pitch height change, while the latency of P3a (an automatic attention shift induced by the detection of deviant features in the passive oddball paradigm) captured the presence of a change in pitch contour (Tsang, Jia, Huang, & Chen, 2011). In Mandarin, pitch height and contour dimensions associated with lexical tone were reported to be lateralized respectively to the right and left hemispheres (Wang, Wang, & Chen, 2013). Their findings, however, may not be attributable to pitch exclusively because standard/deviant tonal contrasts were not phonologically equivalent across experimental conditions. A within-category contrast was used for the height condition; an across-category contrast for the contour condition. The categorical status of tonal contrasts provides a more plausible explanation of the observed pattern of hemispheric laterality (Xi, Zhang, Shu, Zhang, & Li, 2010; Zhang et al., 2011). Though contour and height are important dimensions that are implicated in early, cortical pitch processing, the MMN itself is *not* a pitch-specific response. It is comprised of both auditory and cognitive mechanisms of frequency change detection in auditory cortex (Maess, Jacobsen, Schroger, & Friederici, 2007). This parallel processing is consistent with the near-simultaneity of neurophysiological indicators (EEG/MEG) of psycholinguistic information in the first 200–250 ms (Pulvermuller, Shtyrov, & Hauk, 2009).

The quest to discover an early, preattentive cortical brain response exclusively to pitch began in earnest around the turn of this century. Magnetoencephalography (MEG) was used to study sensitivity to periodicity, an essential requisite of pitch, by investigating the N100m component. However, a large proportion of the N100m is simply a response to the onset of sound energy, and not exclusively to pitch (Alku, Sivonen, Palomaki, & Tiitinen, 2001; Gutschalk, Patterson, Scherg, Uppenkamp, & Rupp, 2004; Hertrich, Mathiak, Lutzenberger, & Ackermann, 2000; Lutkenhoner, Seither-Preisler, & Seither, 2006; Soeta & Nakagawa, 2008; Soeta, Nakagawa, & Matsuoka, 2005; Yrttiaho, Alku, May, & Tiitinen, 2009; Yrttiaho, Tiitinen, Alku, Miettinen, & May, 2010; Yrttiaho, Tiitinen, May, Leino, & Alku, 2008). In order to disentangle the pitch-specific response from the onset response, a novel stimulus paradigm was constructed with two segments – an initial segment of noise with no pitch to evoke the onset components only, followed by a pitch-eliciting segment of iterated rippled noise (IRN) matched in intensity and overall spectral profile (Krumbholz, Patterson, Seither-Preisler, Lammertmann, & Lutkenhoner, 2003). Interestingly, a transient pitch onset response (POR) was evoked from this noise-to-pitch transition only. 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 (Chait, Poeppel, & Simon, 2006; Krumbholz et al., 2003; Ritter, Gunter Dosch, Specht, & Rupp, 2005; Seither-Preisler, Patterson, Krumbholz, Seither, & Lutkenhoner, 2006). 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 (Gutschalk, Patterson, Rupp, Uppenkamp, & Scherg, 2002; Gutschalk et al., 2004; Krumbholz et al., 2003), corroborated by human depth electrode recordings (Griffiths et al., 2010; Schonwiesner & Zatorre, 2008), indicate that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing (Bendor & Wang, 2005; Griffiths, Buchel, Frackowiak, & Patterson, 1998; Johnsrude, Penhune, & Zatorre, 2000; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Penagos, Melcher, & Oxenham, 2004; Zatorre, 1988).

We recently adopted Krumbholz et al.'s (2003) pitch onset response paradigm to demonstrate that a cortical pitch response (CPR) with multiple transient components can be extracted from scalp-recorded electroencephalography (EEG) (Krishnan, Bidelman, Smalt, Ananthkrishnan, & Gandour, 2012). Indeed, neural responses evoked by IRN steady-state pitch stimuli steadily increased in magnitude with increasing IRN stimulus periodicity. Behavioral pitch discrimination also improved with increasing stimulus periodicity. This change in response amplitude with increasing stimulus regularity was strongly correlated with behavioral measures of change in pitch salience. Furthermore, a robust CPR was evoked from both weak and strong IRN pitch-eliciting stimuli, but not to “no-pitch” IRN. We therefore conclude that the CPR is specific to pitch rather than simply a neural response to IRN elicited by slow, spectrotemporal modulations unrelated to pitch (Barker, Plack, & Hall, 2012).

However, any proposed neurobiological mechanism for online processing of pitch contour in the language domain must be able to track dynamic, continuous, nonlinear pitch contours. Besides indexing pitch onset and sound offset, we recently showed that specific components of the CPR mark dynamic pitch attributes of the high rising Tone 2 of Mandarin Chinese (under review). Of the CPR's multiple transient components (Na, Pb, Nb, Pc, Nc), Na–Pb and Pb–Nb showed a systematic increase in interpeak latency and decrease in amplitude with increasing pitch acceleration that followed the time course of pitch change across three within-category variants of Tone 2. Their sensitivity to pitch acceleration was corroborated by strong negative correlations of peak-to-peak amplitude with three measures of pitch acceleration. Na–Pb and Pb–Nb thus appear to be neural markers indexing pitch-relevant neural activity sensitive to the more rapidly-changing portions of the pitch contour. We proposed a series of neural markers embedded in the early stages of cortical sensory processing that flag different temporal attributes of a dynamic pitch contour (Pa: sound onset; Na: pitch onset; Na–Pb/Pb–Nb: pitch change; Pc–Nc: sound offset). We also observed a stimulus-dependent, rightward lateralization at the temporal electrode sites. This hemispheric asymmetry may reflect selective recruitment of experience-dependent, pitch-specific mechanisms in right auditory cortex.

As a logical sequel to our most recent report (under review), the primary objectives of this cross-language study (Chinese, English) are 1) to examine how language experience shapes the

processing of the different temporal attributes of pitch reflected in the CPR components and 2) to determine if the rightward lateralization observed for the Chinese group reflects an experience-dependent functional lateralization of early, sensory-level pitch processing. Our hypothesis is that both the pitch-relevant neural activity indexing the temporal attributes of pitch, and its rightward lateralization at the early sensory level of pitch processing is shaped by language experience. As such, this study is one of a series of CPR experiments that are designed to advance our understanding of early sensory processing of specific temporal attributes of pitch that are present in linguistically-relevant, dynamic pitch contours exemplary of those that occur in natural speech.

2. Materials and methods

2.1. Participants

Ten native speakers of Mandarin Chinese (5 male, 5 female) and English (3 male, 7 female) were recruited from the Purdue University student body to participate in the experiment. 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 closely matched in age (Chinese: 26.0 ± 3.8 years; English: 24.5 ± 3.8), years of formal education (Chinese: 17.4 ± 2.7 years; English: 16.9 ± 1.4), and were strongly right handed (Chinese: $91.2 \pm 15.2\%$; English: 91.0 ± 12.1) as measured by the laterality index of the Edinburgh Handedness Inventory (Oldfield, 1971). All Chinese participants were born and raised in mainland China. None had received formal instruction in English before the age of nine (12.2 ± 1.5 years). As determined by a music history questionnaire (Wong & Perrachione, 2007), all Chinese (except for one) and English participants had less than two years of musical training (Chinese, 1.1 ± 1.2 years; English, 1.1 ± 1.0) on any combination of instruments. No participant had any training within the past five years. Each participant was paid and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

2.2. Stimuli

Three isolated, citation variants of Mandarin Tone 2 were constructed: short (T2_150), intermediate (T2_200), and long (T2_250). Their durations were, in order, 150, 200, and 250 ms. Though infrequent, a short variant (T2_150) has been reported to occur in isolated productions of Tone 2 (Kratochvil, 1985). These durations easily fall outside the range of temporal integration effects (≈ 80 ms) on pitch and its salience for stimuli with resolved harmonics (Plack, Carlyon, & Viemeister, 1995; Plack, Turgeon, Lancaster, Carlyon, & Gockel, 2011; Plack & White, 2000; White & Plack, 1998, 2003). It is therefore unlikely that temporal integration effects pose a potential confound for our evaluation of pitch acceleration-related effects. These stimuli differed in F_0 rate of acceleration as well as duration (Fig. 1). Rates of acceleration, expressed in the acceleration domain, are displayed at 80 ms, minimum-to-maximum, and maximum-to-offset per stimulus (Appendix A.1, table). The maximum speed of pitch change within a speaker's ability to produce a rapid shift in rising pitch over a 4 st interval is 61.3 st/s (Xu & Sun, 2002, p. 1407, Table VII). The average velocity rates (in st/s), calculated from the turning point to F_0 offset, for T2_250 (25.6), T2_200 (32.1), and T2_150 (42.7) fall within the physiological limits of speed of rising pitch changes. As reflected by FFR responses in the brainstem (Krishnan, Gandour, Smalt, & Bidelman, 2010, p. 96, Figs. 2 and 3), a scaled variant of Tone 2 with a velocity rate of 51.94 st/s, though approaching the upper bounds of the normal voice range, was statistically indistinguishable from an exemplary Tone 2 stimulus (25.4 st/s). To enable us to focus primarily on the effects of changes in rate of acceleration during the rising portion of Tone 2, F_0 onset (100.88 Hz) and offset (131.72 Hz) were held constant across stimuli. ΔF_0 from turning point to offset was fixed across stimuli at 30.84 Hz (4.6 st; .38 octaves). This ΔF_0 value is comparable to that of an exemplary Tone 2 citation form (Krishnan, Gandour, Smalt, et al., 2010) and is an effective cue for the perception of isolated Tone 2 (Moore & Jongman, 1997). The turning point was located at about $\approx 26\%$ of the duration of the F_0 contour (40 ms, T2_150; 53 ms, T2_200; 66 ms, T2_250). The timing of these turning points relative to F_0 onset are perceptually relevant in the identification of Tone 2 (cf. Moore & Jongman, 1997,

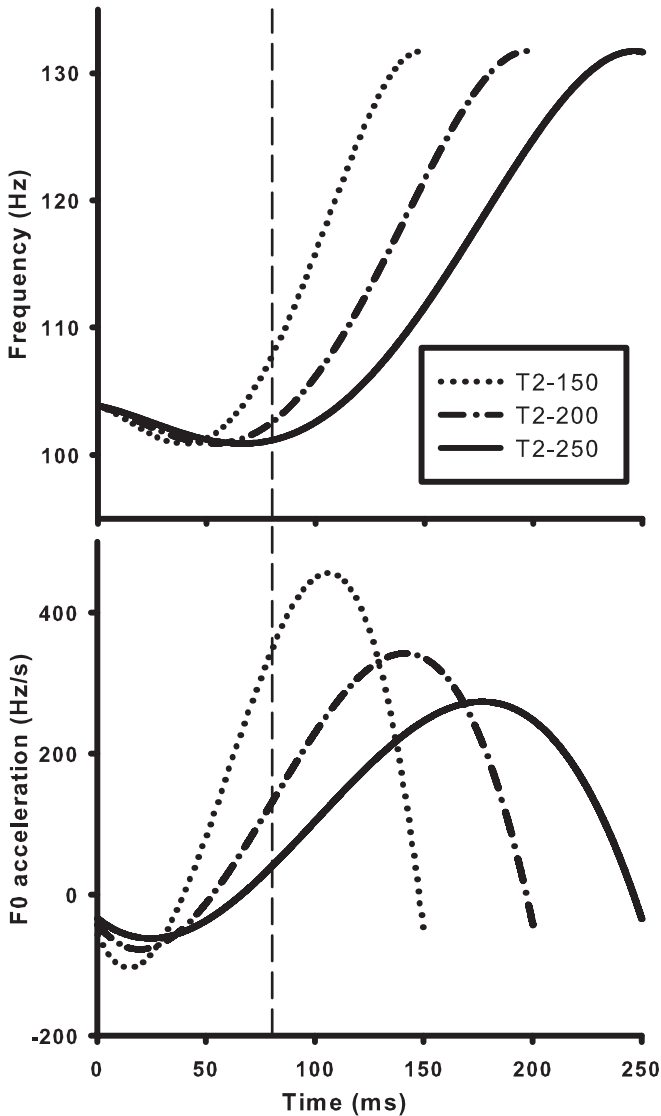


Fig. 1. IRN stimuli used to evoke cortical responses to linguistic patterns that are differentiated by varying degrees of rising acceleration and duration. Voice fundamental frequency (F_0) contours (top panel) and corresponding acceleration trajectories (bottom panel) of all three stimuli are modeled after the citation form of Mandarin Tone 2 (T2) using a fourth-order polynomial equation. These three stimuli exemplify short (T2_150, dot), intermediate (T2_200, dash-dot), and long (T2_250, solid) variants of Tone 2. The vertical dashed line at 80 ms is located after the turning point, and provides a measure of instantaneous acceleration irrespective of stimulus duration (53, 40, and 32% of total duration for T2_150, T2_200, and T2_250, respectively).

p. 1870, Fig. 4). Based on these behavioral and neural data, we judged these stimuli to be ecologically valid (within-category) representations of Tone 2 and likely to elicit differential sensitivity to varying degrees of acceleration rates at the cortical level.

These three IRN stimuli with time-varying F_0 contours were generated by applying a *time-varying*, delay-and-add algorithm using fourth-order polynomial equations (Appendix A.2, text) (Denham, 2005; Krishnan, Swaminathan, & Gandour, 2009; Sayles & Winter, 2007; Swaminathan, Krishnan, Gandour, & Xu, 2008). A high iteration step ($n = 32$) was chosen because pitch salience does not

increase by any noticeable amount beyond this number of iteration steps. The gain was set to 1. By using IRN, we preserve dynamic variations in pitch of auditory stimuli that lack a waveform periodicity, formant structure, temporal envelope, and recognizable timbre characteristic of speech.

Each stimulus condition consisted of two segments (crossfaded with 5 ms cos² ramps): an initial 500 ms noise segment followed by a pitch segment, i.e., T2_150, T2_200, and T2_250 (Appendix B.1, figure). The overall RMS level of each segment was equated such that there was no discernible difference in intensity between initial and final segments. All stimuli were presented binaurally at 80 dB SPL through magnetically-shielded tubal insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA) with a fixed onset polarity (rarefaction) and a repetition rate of .94/s. Stimulus presentation order was randomized both within and across participants. All stimuli were generated and played out using an auditory evoked potential system (SmartEP, Intelligent Hearing Systems; Miami, FL, USA).

2.3. Cortical pitch response 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), ignore the sounds they heard, and were allowed to sleep throughout the duration of the recording procedure ($\approx 75\%$ fell asleep). The EEG was acquired continuously (5000 Hz sampling rate; 3–2500 Hz analog band-pass) using ASA-Lab EEG system (ANT Inc., The Netherlands) utilizing a 32-channel amplifier (REFA8-32, TMS International BV) and WaveGuard (ANT Inc., The Netherlands) electrode cap with 32-shielded sintered Ag/AgCl electrodes configured in the standard 10–20-montage system. The high sampling rate of 5 kHz was necessary to recover the brainstem frequency following responses in addition to the relatively slower cortical pitch components. Because the primary objective of this study was to evaluate the effects of language experience on the characteristics of cortical pitch components, and not their source localization, EEG acquisition was accomplished using an electrode montage including the following 9 electrode locations: Fpz, AFz, Fz, F3, F4, Cz, T7, T8, M1, M2 (Appendix B.2, figure). 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 Ω . For each stimulus, EEGs were acquired in blocks of 1000 sweeps. The experimental protocol took about 2 h to complete.

2.4. Extraction of the cortical pitch response (CPR)

CPR responses were extracted off-line from the EEG files. To extract the cortical pitch response components, EEG files were first down sampled from 5000 Hz to 2048 Hz. They were then digitally high-pass filtered (3–25 Hz) to enhance the transient components and minimize the sustained component. Sweeps containing electrical activity exceeding $\pm 40 \mu\text{V}$ were rejected as artifacts. 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 laterality effects. The re-referenced electrode site, Fz-linked T7/T8, was used to characterize the transient pitch response components. For both averaging procedures, the analysis epoch was 1200 ms including the 100 ms pre-stimulus baseline.

2.5. Analysis of CPR

The CPR is characterized by obligatory components (P1/N1) corresponding to the onset of energy in the precursor noise segment of the stimulus followed by an onset component (Pa) and four transient, pitch-related response components (Na, Pb, Nb, Pc) occurring after the onset of the pitch-eliciting segment of the stimulus. To characterize what aspects of the dynamic pitch contours are being

indexed by the components of the CPR, the latency and magnitude of only the CPR components were evaluated. Peak latencies of response components corresponded to the time interval between the pitch-eliciting stimulus onset and a response peak of interest: Pa, Na, Pb, Nb, and Pc. Their interpeak latencies corresponded to the time interval between adjacent response peaks of these components: Na–Pb, Pb–Nb, and Nb–Pc. These latency measures enabled us to identify the components associated with pitch onset, pitch acceleration, and stimulus offset. Pa–Pc interpeak latency was measured to identify an interval that marks stimulus duration. Peak-to-peak amplitude of Na–Pb, Pb–Nb, and Nb–Pc was measured to determine if variations in amplitude were indexing specific aspects of the pitch contour (pitch onset, changes in pitch, pitch offset). In addition, peak-to-peak amplitude of Na–Pb and Pb–Nb was measured separately at the frontal (F3/F4) and temporal (T7/T8) electrode sites to evaluate laterality effects. To enhance visualization of the laterality effects along a spectrotemporal dimension, joint time-frequency analysis was performed using a continuous wavelet transform on the grand average waveforms derived from the frontal and temporal electrodes.

2.6. Statistical analysis

Separate three-way, mixed model ANOVAs (SAS[®]; SAS Institute, Inc., Cary, NC, USA) were conducted on peak latency, interpeak latency, and peak-to-peak amplitude derived from the Fz electrode site. Language group (Chinese, English) was treated as a between-subjects factor; subjects as a random factor nested within group. Stimulus (T2_150, T2_200, T2_250) and component were treated as within-subject factors. In the analysis of peak latency, there were four components (Na, Pb, Nb, Pc); in the analysis of interpeak latency and peak-to-peak amplitude, three components (Na–Pb, Pb–Nb, Nb–Pc). By analyzing these components, we were able to assess the effects of pitch acceleration on latency and amplitude across stimuli. Separate two-way (group \times hemisphere), mixed model ANOVAs were conducted on peak-to-peak amplitude of Na–Pb and Pb–Nb derived from the T7/T8 (temporal) and F3/F4 (frontal) electrode sites for T2_250 only. By focusing on these two pitch-related components, we were able to determine whether laterality effects at the frontal and temporal sites vary as a function of language experience. T2_250 was chosen because it was the only one to show an RH (T8 > T7) advantage in peak-to-peak amplitude for the native Chinese group (Krishnan, Gandour, Ananthakrishnan, & Vijayaraghavan, 2014). All a priori or post hoc multiple comparisons were corrected with a Bonferroni adjustment at $\alpha = 0.05$.

3. Results

3.1. Response morphology of CPR components

Grand averaged cortical evoked response waveforms to the three stimuli are shown for the Chinese (red trace) and the English (blue trace) group in Fig. 2. The top panel shows both the superimposed P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250. As expected (Krishnan, Bidelman, et al., 2012) the obligatory P1/N1 complex, reflecting neural activity synchronized to the onset of the noise precursor (black), is very similar for both groups, and across the three stimulus conditions. The CPR components, elicited by the pitch-eliciting stimulus segment, are characterized by a series of successive biphasic components (in ms): e.g., T2_250 (bottom panel), Pa, 70–85; Na, 125–150; Pb, 200–220; Nb, 270–285; Pc, 305–325; and Nc, 340–360. The second, third, and fourth (bottom) panels show only the CPR waveforms elicited by the three stimuli. The CPR components are clearly identifiable for both groups. The amplitude of components Na, Pb, and Nb for T2_200 and T2_250, however, are greater for the Chinese group. The increase in amplitude for these components from T2_150 to T2_250 is more apparent for the Chinese group. The offset components (Pc, Nc) appear to be more robust for the English group. Peak latency for Pa and Na do not change appreciably across stimulus conditions for either group. In contrast, response components Pb, Nb, Pc, and Nc all show a systematic increase in peak latency across stimulus conditions for both groups. Consistent with these observations, the interpeak latencies (Na–Pb, Pb–Nb, Nb–Pc) increase across stimulus conditions. These systematic changes in response amplitude and latency are likely produced by a decrease in the rate of pitch acceleration

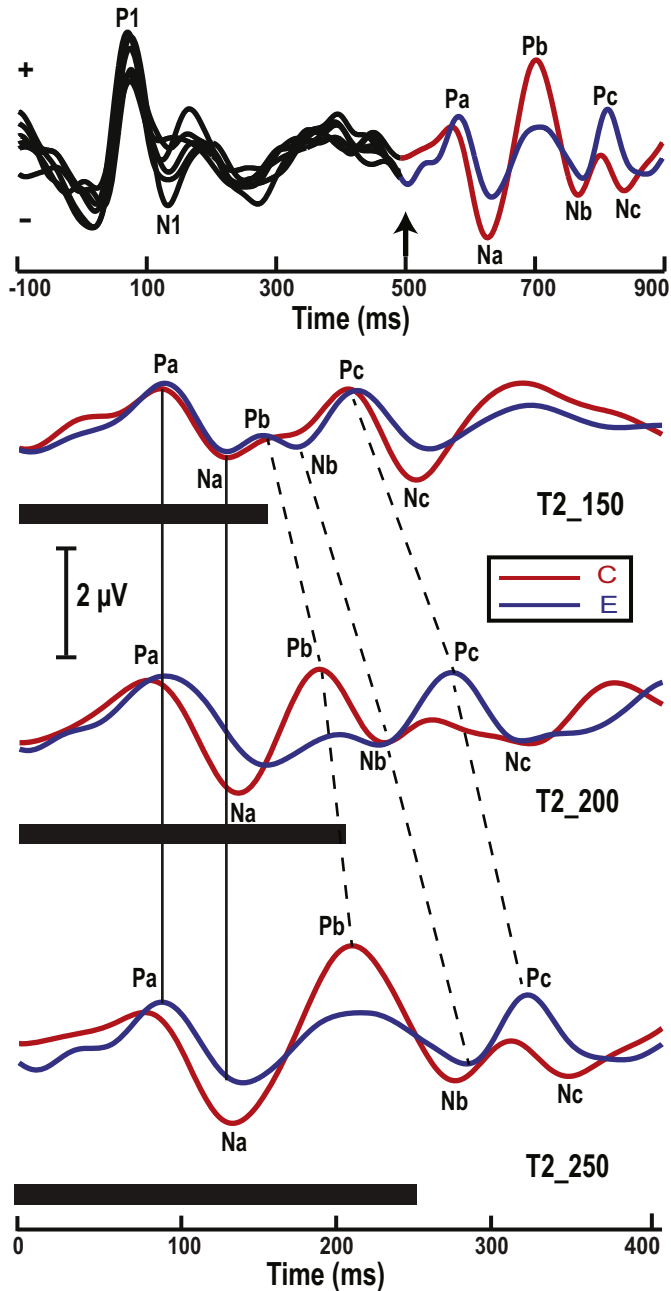


Fig. 2. Grand averaged cortical evoked response components recorded at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (Chinese, red; English, blue) are displayed in the top panel. The up arrow at 500 ms marks the onset 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.)

and an increase in duration across the three stimulus conditions. The more robust response amplitude of the CPR components in the Chinese group may reflect an experience-dependent enhancement of components related to pitch.

3.2. Latency and amplitude of CPR components

3.2.1. Latency

For both language groups, mean peak latencies of Pb, Nb, and Pc components increase systematically regardless of stimulus (T2_150, T2_200, T2_250) with appreciably smaller increases for Na (Appendix B.3, figure). ANOVA results showed a three-way interaction among group, component, and stimulus ($F_{6,108} = 3.59, p = 0.0028$). Planned group comparisons indicated that Chinese latencies were shorter than English for Na and Pb in response to T2_200, and for Pc in response to T2_250. Other group comparisons failed to reach significance.

Mean interpeak latency for Na–Pb, Pb–Nb, and Nb–Pc components generally show a systematic increase from T2_150 to T2_250 for Na–Pb and Pb–Nb in both language groups (Fig. 3a). ANOVA results revealed a three-way interaction among group, component, and stimulus ($F_{6,108} = 3.95, p = 0.0013$). Planned group comparisons indicated that in response to T2_200, Chinese interpeak latency was longer

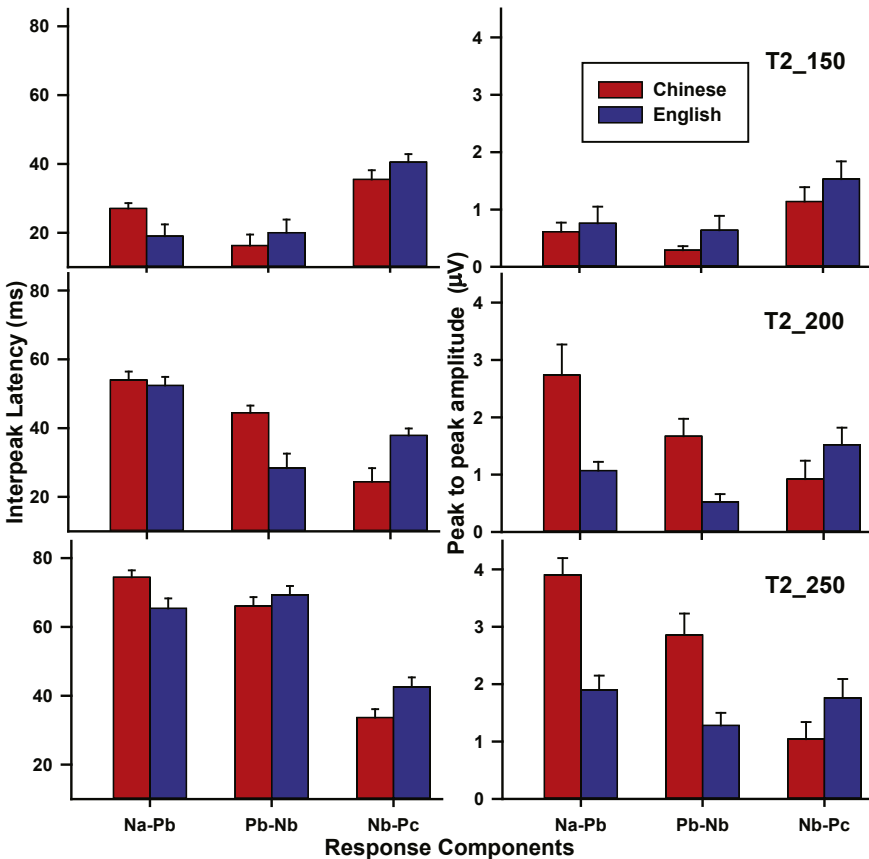


Fig. 3. Mean interpeak latency (left panel) and peak-to-peak amplitude (right panel) of Na–Pb, Pb–Nb, and Nb–Pc components recorded at the Fz electrode site from T2_150 (top panel) to T2_250 (bottom panel) in both Chinese and English groups. Interpeak latencies increase across stimuli for Na–Pb and Pb–Nb 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 a larger amplitude than the English group for Na–Pb and Pb–Nb in T2_200 and T2_250. Error bars = ± 1 SE.

than English for Pb–Nb, but shorter than English for Nb–Pc. Other group comparisons failed to reach significance.

Pa–Pc, the component that closely corresponded to stimulus duration, exhibited longer interpeak latencies as one progresses from T2_150 to T2_250 (Appendix A.3, table). A two-way ANOVA of group and stimulus showed only a main effect of stimulus ($F_{2,36} = 1295.99, p < 0.0001$), meaning that Chinese and English listeners were homogenous with respect to this duration-related component.

3.2.2. Amplitude

Mean peak-to-peak amplitude of Na–Pb, Pb–Nb, and Nb–Pc components show that Chinese listeners exhibit a larger amplitude than English for Na–Pb and Pb–Nb in T2_200 and T2_250 (Fig. 3b). ANOVA results revealed a three-way interaction among group, component, and stimulus ($F_{6,108} = 3.04, p = 0.0087$). In response to T2_250, planned group comparisons indicated that Chinese peak-to-peak amplitude was larger than English for both Na–Pb and Pb–Nb. In response to T2_200, Chinese peak-to-peak amplitude was larger than English for Na–Pb; for Pb–Nb, the Chinese advantage was marginally larger than English ($p = 0.0797$). Other group comparisons failed to reach significance.

To support the view that Chinese superiority on peak-to-peak amplitude reflects enhanced neural activity associated with rapidly-changing pitch, Pearson correlations (r) were computed between peak-to-peak amplitude of CPR components (Na–Pb, Pb–Nb) and three measures of pitch acceleration for T2_250 (Appendix A.1, table). In the Chinese group, a strong negative association was observed between Na–Pb ($r = -.781/-.779$) and Pb–Nb ($r = -.774/-.764$) with all measures of pitch acceleration (Appendix A.4, table). In the English group, we observed a much weaker negative association (Na–Pb, $r = -.519/-.497$; Pb–Nb, $r = -.322/-.290$). The negative correlation coefficient means that peak-to-peak amplitude increases with decreasing acceleration.

A discriminant analysis was used to determine the extent to which individual participants may be classified into their respective language groups based on their peak-to-peak magnitude values for T2_250. Overall, 95% of participants were correctly classified into their respective language groups (Chinese, 90%; English, 100%) (Appendix A.5, table). Because we can expect to get only 50% of the classifications correct by chance, an overall accuracy rate of 95% represents a considerable improvement (canonical correlation = .796). Only .05% fewer correct classifications (Chinese, 9/10; English, 9/10) were made in the cross-validated analysis in comparison to the original analysis. The group centroids, i.e., average discriminant z scores, differed significantly between the Chinese (1.248) and English (–1.248) groups ($F_{2,17} = 14.71, p = 0.0002$). The pooled within-class standardized canonical coefficients for Na–Pb and Pb–Nb, respectively, were .947 and .379, indicating that Na–Pb was more than twice as important as Pb–Nb in discriminating listeners by language affiliation.

3.3. Comparison of CPR components at frontal (F3/F4) and temporal (T7/T8) electrode sites to examine hemispheric laterality

The grand average waveforms (two left columns) and their corresponding spectra (two right columns) of the CPR components for each of the three stimuli per language group are displayed at frontal (F3/F4) and temporal (T7/T8) sites in Figs. 4 and 5, respectively. The waveform data in Fig. 4 reveal that pitch-related components at frontal F3 (dashed waveforms) and F4 (solid waveforms) are more robust for the Chinese group (red waveforms in the first column) across all three stimuli when compared to the smaller CPR components for the English group (blue waveforms in the second column). However, for both groups, the CPR components at the two electrode sites essentially overlap with no discernible difference in magnitude and thus no laterality of the CPR components. The lack of laterality at these electrode sites is evident in their essentially identical spectral plots (two right columns). Similarly, CPR components at both T7 and T8 electrode sites (Fig. 5) are larger in amplitude for the Chinese group (red waveforms, in the web version) relative to the English group (blue waveforms, in the web version) particularly for T2_200 and T2_250. In contrast to the F3/F4 waveforms, these same components are larger at the right temporal electrode (T8: solid red, in the web version) than the left temporal electrode (T7: dashed red, in the web version) for the Chinese group exclusively and for T2_250 in particular. This robust rightward lateralization for T2_250 is clearly evident in the spectrotemporal representation of the pitch-related components (bottom two right panels).

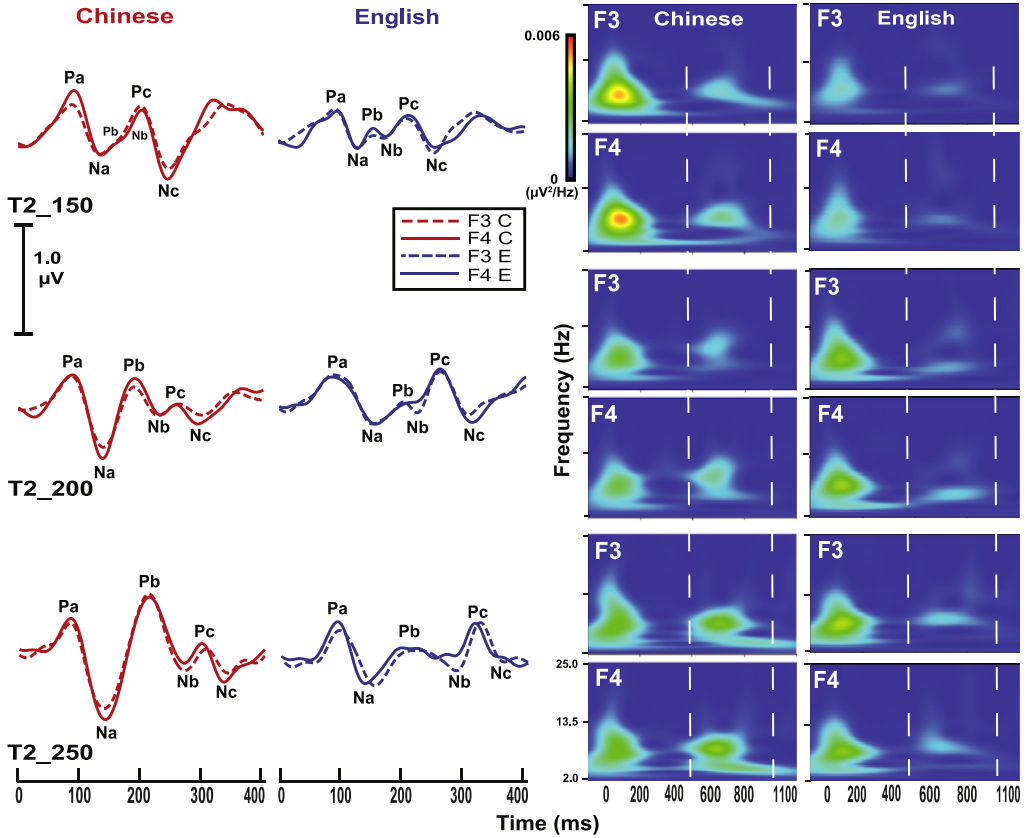


Fig. 4. Grand average waveforms (two left columns) and their corresponding spectra (two right columns) of the CPR components for the two groups (red: Chinese; blue: English) recorded at electrode sites F3 (dashed lines) and F4 (solid lines) for each of the three stimuli. The zero on the x-axis denotes the time of onset of the pitch-eliciting segment of the three stimuli. The response components are generally greater in magnitude for the Chinese group compared to the English group with no discernible asymmetry between F3 and F4 for either group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Mean peak-to-peak amplitude of Na–Pb and Pb–Nb components for the Chinese and English groups are displayed in response to T2_250 at temporal (T7/T8) and frontal (F3/F4) sites in Fig. 6. At the T7/T8 electrode sites, ANOVA results irrespective of component showed only main effects of group, Chinese > English (Na–Pb: $F_{1,18} = 27.51$, $p < 0.0001$; Pb–Nb, $F_{1,18} = 23.12$, $p < 0.0001$) and hemisphere, T8 > T7 (Na–Pb: $F_{1,18} = 7.72$, $p = 0.0124$; Pb–Nb, $F_{1,18} = 6.31$, $p = 0.0217$). The interaction effect (group \times hemisphere) failed to reach significance for either component, meaning that peak-to-peak amplitude was larger in the Chinese group as compared to the English group across temporal electrode sites; and that peak-to-peak amplitude was larger in the right than in the left temporal site regardless of group. In contrast, at the F3/F4 electrode sites, ANOVA results showed only a main effect of group, Chinese > English, regardless of component (Na–Pb: $F_{1,18} = 20.56$, $p = 0.0003$; Pb–Nb, $F_{1,18} = 44.22$, $p < 0.0001$). The hemisphere main effect failed to reach significance, meaning that peak-to-peak amplitude of these CPR components did not vary between the left and right frontal sites. The absence of an interaction effect means that Chinese peak-to-peak amplitude was larger than English at either frontal site.

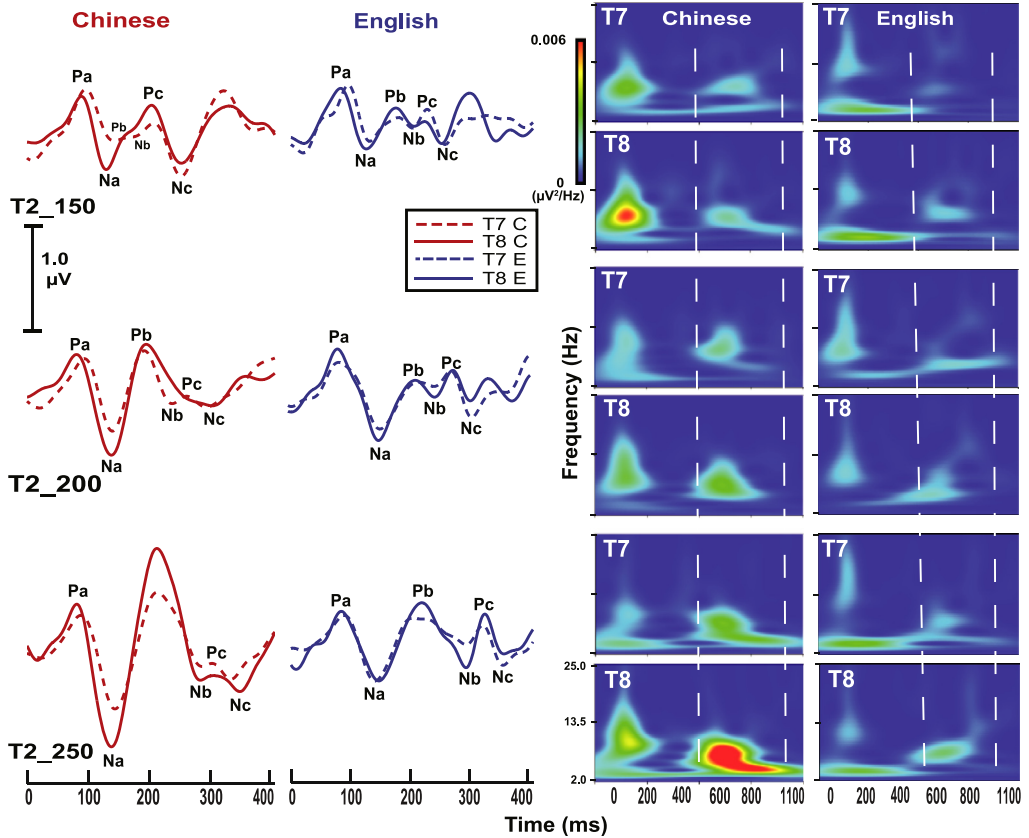


Fig. 5. Grand average waveforms and their corresponding spectra of the CPR components for the two groups recorded at electrode sites T7 and T8 for each of the three stimuli. The response components are generally greater in magnitude for the Chinese group compared to the English group with a large rightward asymmetry for the Chinese group only for stimulus T2_250. See also caption to Fig. 4.

4. Discussion

The major findings of this cross-language study show that the magnitude of CPR components (Na–Pb and Pb–Nb) and the correlation between these two components and pitch acceleration are stronger for the Chinese listeners compared to English listeners for stimuli that fall within the range of a native pitch contour (as produced on isolated monosyllables). Taken together, these findings suggest that long-term language experience shapes early sensory level processing of pitch in the auditory cortex. The sensitivity of the CPR may vary depending on the relative linguistic importance of specific temporal attributes of dynamic pitch. As revealed by discriminant function analysis, the Na–Pb component was more than twice as important as Pb–Nb in grouping listeners by language affiliation. A stronger rightward asymmetry at the temporal electrode sites for Chinese listeners, relative to English listeners, is compatible with the notion of experience-dependent modulation of pitch-specific mechanisms at an early stage of processing in right auditory cortex.

4.1. Experience-dependent neural plasticity in early sensory processing of pitch in the auditory cortex

Our findings are consistent with earlier cross-language studies that have revealed experience-dependent neural plasticity at both cortical and subcortical levels of the brain (Gandour &

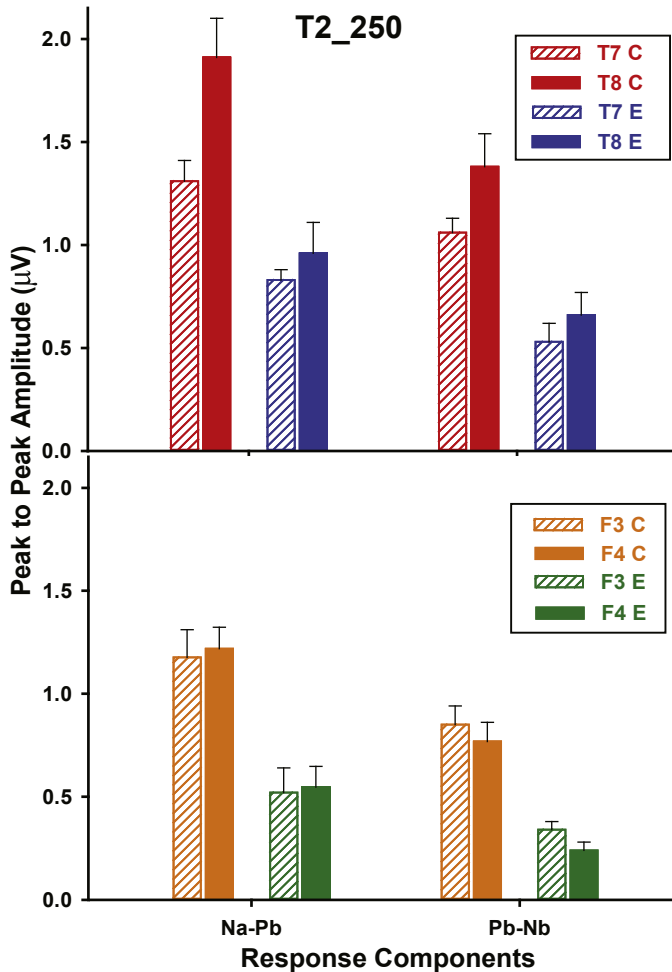


Fig. 6. Mean peak-to-peak amplitude of Na–Pb and Pb–Nb components for the Chinese and English groups in response to T2_250 at temporal (T7/T8; top panel) and frontal (F3/F4; bottom panel) sites. At the T7/T8 electrode sites, peak-to-peak amplitude is larger in the Chinese group than the English group in both hemispheres. A right-sided advantage is present in each language group. However, this rightward asymmetry is more robust in the Chinese group compared to the relatively weak asymmetry in the English group. At the F3/F4 electrode sites, Chinese peak-to-peak amplitude is larger than English in both hemispheres, though there is no hemispheric advantage for either language group. Error bars = ± 1 SE.

Krishnan, 2014; Krishnan et al., 2014; Krishnan, Gandour, et al., 2012; Zatorre & Baum, 2012; Zatorre & Gandour, 2008). We believe that long-term experience-driven adaptive pitch mechanisms at early sensory levels of pitch processing in the auditory cortex sharpen response properties of neural elements to enable optimal representation of temporal attributes of native pitch contours. In this study, all three stimuli represented variant productions of Mandarin Tone 2, though T2_150 was marginal as spoken on isolated monosyllables (Kratochvíl, 1985). A language-dependent effect on peak-to-peak amplitude was observed for T2_250 and T2_200 only. Thus, not all within-category representations of a tonal category are equal in terms of their influence on early cortical pitch processing.

We recently reported a systematic increase in the interpeak latency and decrease in amplitude for components Na–Pb and Pb–Nb with increasing pitch acceleration (Krishnan et al., 2014). We inferred

that these components may be indexing pitch-relevant neural activity associated with the more rapidly-changing portions of the pitch contour. This inference was further strengthened by a strong correlation with pitch acceleration for Na–Pb and Pb–Nb only. On these same components, the Chinese group exhibited greater amplitude and higher correlation with pitch acceleration than the English group. This language-dependent effect suggests an experience-dependent increase in sensitivity to dynamic portions of pitch contours that occur in the native listeners' experience. Because enhanced sensitivity to time-varying dimensions (e.g., acceleration) is already present in pitch encoding at the level of the brainstem (Krishnan & Gandour, 2009; Krishnan, Gandour, et al., 2012), it seems plausible that cortical pitch mechanisms may be reflecting, at least in part, this enhanced pitch input from the brainstem.

Our current experimental design does not permit us to determine whether Na–Pb and Pb–Nb index different dynamic segments of the pitch contour. We hypothesize that Na–Pb (relatively longer latency and larger amplitude) indexes the increasing pitch acceleration between the turning point and the point of maximum acceleration in the stimulus. Whereas Pb–Nb (shorter latency and smaller amplitude) indexes the shorter pitch deceleration between maximum acceleration and stimulus offset. Interestingly, discriminant analysis showed that Na–Pb contributed more to the accurate grouping of listeners by language affiliation.

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 event markers of neural activity indexing the temporal course of a pitch contour. By design, this experiment minimized latency effects. Pitch height was fixed; timing differences from onset to turning point across stimuli were very small (in ms: T2_150, 40 T2_200, 53; T2_250, 66); the turning point itself occurred at about 26% of total duration across stimuli. Future research is clearly warranted to investigate how the latency of CPR components is exploited to signal specific temporal attributes of pitch contour (cf. Tsang et al., 2011).

4.2. Influence of language experience on the hemispheric preferences for pitch processing

We observed a stronger rightward asymmetry of Na–Pb and Nb–Pb responses by the Chinese group, relative to the English group. This language-dependent effect was confined to T2_250 only. Of the three stimuli, T2_250 most closely approximates the canonical duration pattern of Tone 2 produced in isolation ($M = 273$ ms: Xu, 1997). The English group also displayed a rightward asymmetry albeit much weaker than the Chinese group. One possible explanation involves the distinction between the *sensory memory trace* and *analyzed sensory memory* (Cowan, 1984, 1987). The latter contains fine-grained, analyzed sensory codes including time-varying (e.g., pitch slope or acceleration) and event-timing (e.g., onset time or duration) information. Its lifetime is on the order of seconds. Why an experience-dependent effect occurs only on the stimulus that best exemplifies the tonal category requires explanation. Bear in mind that the experiment is free of task demands; stimuli are reduced to the pitch parameter; and hemispheric asymmetry is based on peak-to-peak amplitude responses extracted from two putative, pitch-specific components (Na–Pb, Pb–Nb). This differential sensitivity to within-category representations leads us to hypothesize that pitch information is encoded in a hierarchical order including a *short-term categorical memory* that interacts with analyzed sensory memory within the same time interval (cf. Goldinger, 1998; Pasternak & Greenlee, 2005; Xu, Gandour, & Francis, 2006). The English group obviously would have no memory of the canonical pattern of Mandarin Tone 2. The asymmetry was confined to the temporal electrodes (T8 > T7). No asymmetry was found at the frontal electrodes sites (F3/F4) regardless of stimulus or language group. The fact that hemispheric asymmetry occurred in auditory areas, but not frontal, suggests that different mechanisms and networks are involved at lower-level stages of pitch processing. That dorsal regions of the right superior temporal gyrus play a critical role in early stages of processing suprasegmental information is incontrovertible (Friederici & Alter, 2004; Meyer, 2008; Zatorre & Gandour, 2008). However, less is known about the nature of the interaction

between the right auditory core and adjoining auditory-related cortical areas. One view is that auditory processing occurs symmetrically in the core, but asymmetrically in auditory-related areas (Poeppel, 2003; Poeppel, Idsardi, & van Wassenhove, 2008). In this study, we hypothesize that the language-dependent temporal asymmetry in 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 interaction between sensory and cognitive components within the language domain in *right* auditory-related cortex. Indeed, a complete account of pitch processing must allow for interactions between sensory and cognitive contributions that interact *within the same time interval*, as well as at different time intervals at different cortical levels within and across hemispheres.

Our finding of stronger rightward asymmetry of pitch-relevant neural activity for the Chinese listeners converges with ERP data that reveal the emergence of experience-dependent asymmetries in the music domain at early cortical levels of processing. For example, a right temporal advantage is seen in the cortical N1 component related to *pitch transition* (change-N1, ~100 ms latency) in trained musicians (Itoh, Okumiya-Kanke, Nakayama, Kwee, & Nakada, 2012). No hemispheric asymmetry is observed for the onset component. Using musical pitch stimuli, the Itoh et al. study similarly demonstrates experience-dependent enhancement of processing changes in pitch in the right auditory cortex.

We must also point out that our stimuli exhibit dynamic, *curvilinear* F_0 trajectories that are representative of a Mandarin lexical tone. Steady-state or flat F_0 patterns are of no functional relevance in the speech of any of the world's languages, tonal or otherwise. Interestingly, MEG recordings fail to observe any hemispheric differences with regard to either latency or amplitude of the pitch-relevant cortical components elicited by stimuli with *flat* pitch (Gutschalk et al., 2004; Hari et al., 1987; Krumbholz et al., 2003; Lutkenhoner & Steinstrater, 1998; Seither-Preisler et al., 2006). This disparity in hemispheric asymmetry between dynamic and flat pitch patterns further emphasizes the importance of using ecologically-relevant stimuli to study pitch processing in the language domain.

4.3. Cross-language differences in relative importance of pitch attributes as reflected by CPR components

The discriminant function analysis of Fz peak-to-peak amplitude (elicited by T2_250) was highly successful in separating the two language groups. The relative weighting of CPR components showed that Na–Pb is twice as important as Pb–Nb in classifying Chinese and English listeners into their respective group. While our current experimental design does not permit us to determine whether components Na–Pb and Pb–Nb are indexing different portions of the dynamic segment of the pitch contour, we hypothesize that Na–Pb (relatively longer latency and larger amplitude) indexes the rapid increase in pitch acceleration between the turning point and the point of maximum acceleration in the stimulus; whereas Pb–Nb (shorter latency and smaller amplitude) indexes the shorter pitch deceleration segment between maximum acceleration and stimulus offset. Because rapid changes in pitch at the syllable level is one of the critical features of a contour-tone language (Pike, 1948), native speakers of Mandarin place more emphasis on Na–Pb, relative to non-tone language speakers, in early cortical stages of pitch extraction from the auditory signal. These language-dependent effects are manifest even though the electrophysiological responses themselves are pitch-specific. That is, language experience may influence electrophysiological responses to *temporal attributes of pitch* rather than holistic, tonal categories. This is not surprising if one adopts a parallel model of brain processing. It is well-known that early, near-simultaneous brain indexes of a range of psycholinguistic processes emerge within 100–250 ms after critical stimulus information is present (Pulvermuller et al., 2009). Moreover, CPR components permit us to investigate the *dynamic* portion of a lexical tone, which may lead to a fuller understanding of real-time neurobiological mechanisms that follow the time course of a pitch contour. And finally, these CPR data extend our previous findings on the relative weighting of dimensions or attributes of pitch at the levels of the cerebral cortex (MMN: Chandrasekaran et al., 2007) and brainstem (FFR: Krishnan, Gandour, & Bidelman, 2010; Krishnan et al., 2009).

4.4. Neural mechanisms mediating experience-dependent plasticity of early sensory processing of pitch in the 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 (Gilbert & Sigman, 2007).

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 shifting of best frequencies toward task-relevant, target stimuli (Fritz, Shamma, Elhilali, & Klein, 2003; Lee & Middlebrooks, 2011; see Weinberger, 2011, for review); and selective expansion of receptive fields for stimulus features that are being learned (Polley, Steinberg, & Merzenich, 2006). 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 (Russo, Nicol, Zecker, Hayes, & Kraus, 2005; Song, Skoe, Wong, & Kraus, 2008); long-term linguistic experience (Krishnan & Gandour, 2009; Krishnan, Gandour, et al., 2012; Krishnan, Xu, Gandour, & Cariani, 2005); and musical training (Bidelman & Krishnan, 2009; Bidelman, Krishnan, & Gandour, 2011; Musacchia, Sams, Skoe, & Kraus, 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007).

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 (Ahissar & Hochstein, 2004; Nahum, Nelken, & Ahissar, 2008). 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 top-down influences on cortical (Krizman, Skoe, Marian, & Kraus, 2014) and subcortical sensory processing (Banai, Abrams, & Kraus, 2007; Krishnan, Bidelman, & Gandour, 2010). 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, spatio-temporally broader, pitch-relevant information in the auditory cortex (Chechik et al., 2006; Warren & Griffiths, 2003; Winer, Miller, Lee, & Schreiner, 2005; Zatorre & Belin, 2001). 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 (Krishnan, Bidelman, et al., 2012).

Another proposed circuitry mediating learning-induced plasticity is the cortico-colliculo-thalamo-cortico-collicular loop (Xiong, Zhang, & Yan, 2009). This circuitry is comprised of bottom-up (colliculo-thalamic and thalamo-cortical) and top-down (corticofugal) projections that form a tonotopic loop. It is presumed to be the only neural substrate that carries accurate auditory information (cf. Krishnan & Gandour, 2009). Additionally, 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 (Bajo, Nodal, Moore, & King, 2010). 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 (Krishnan, Bidelman, et al., 2012). In the case of humans, top-down processes likely shape the reorganization of the sensory processing of pitch-relevant information in the brainstem and auditory cortex to enhance pitch extraction in earlier stages of language development when adaptive plasticity presumably would be most vigorous (Keuroghlian & Knudsen, 2007; Kral & Eggermont, 2007). The slower time constants of corticofugal processing render it much too sluggish to effectively influence a dynamic pitch pattern over its entire duration (Dean, Robinson, Harper, & McAlpine, 2008). Nonetheless, its adaptive properties would still be able to facilitate extraction of behaviorally-relevant information under degraded listening conditions and during training protocols.

4.5. *Neural mechanism(s) for early sensory level pitch processing in the auditory cortex*

It is generally agreed that lateral Heschl's gyrus is the putative source for the pitch onset component (Na). Generator sources for the remaining pitch-relevant components (Pb, Nb) are unknown and cannot be determined from this study. 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 (Cariani & Delgutte, 1996a, 1996b; Cedolin & Delgutte, 2005; Plack et al., 2005). There is evidence that neurons in primary auditory cortex exhibit temporal and spectral response properties that could enable these pitch-encoding schemes (Lu, Liang, & Wang, 2001; Steinschneider, Reser, Fishman, Schroeder, & Arezzo, 1998), 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 F_0 . 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 (Walker, Bizley, King, & Schnupp, 2011). 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.

It has been proposed that processing of specific pitch values, pitch salience and pitch change occurs in the lateral Heschl's gyrus well after the time-interval processing begins in subcortical regions to encode pitch relevant information (Griffiths, Uppenkamp, Johnsrude, Josephs, & Patterson, 2001; Langner & Schreiner, 1988; Patterson et al., 2002; Winter, Wiegrebe, & Patterson, 2001). Gutschalk et al. (2004) have further suggested that the cortical pitch response more likely represents the integration of pitch information across frequency channels and/or the calculation of specific pitch value and pitch strength in Heschl's gyrus. This is because the latency of the cortical pitch response is too long to represent the temporal processing required to generate the auditory image response in the subcortical structures.

4.6. *Conclusions*

Our discovery of cortical pitch components that index several behaviorally-relevant temporal attributes of dynamic, curvilinear pitch contours that are ecologically representative of natural speech provides a new avenue to evaluate pitch processing at different levels of the brain. Both stimulus-dependent enhancement and stronger rightward asymmetry of CPR components in the Chinese group is consistent with the notion that early sensory-level pitch processing in the auditory cortex is shaped by language experience. This long-term 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. With this novel technique, we now have a physiologic window to evaluate the interplay between bottom-up,

top-down, and local intrinsic components in the hierarchical processing of pitch-relevant information (cf. Foxe & Schroeder, 2005).

Acknowledgments

Research supported by NIH 5R01DC008549-07 (A.K.). Thanks to Longjie Cheng for her assistance with statistical analysis (Department of Statistics); Jillian Wendel and Chandan Hunsur Sarresh 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, USA, or via email: rkrish@purdue.edu.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jneuroling.2014.08.002>.

References

- 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.
- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8(10), 457–464. <http://dx.doi.org/10.1016/j.tics.2004.08.011>. S1364-6613(04)00215-3 [pii].
- Alku, P., Sivonen, P., Palomaki, K., & Tiitinen, H. (2001). The periodic structure of vowel sounds is reflected in human electro-magnetic brain responses. *Neuroscience Letters*, 298(1), 25–28.
- 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. <http://dx.doi.org/10.1038/nn.2466>.
- 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. <http://dx.doi.org/10.1080/14992020701383035>, 781872134 [pii].
- Barker, D., Plack, C. J., & Hall, D. A. (2012). Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. *Cerebral Cortex*, 22(4), 745–753. <http://dx.doi.org/10.1093/cercor/bhr065>.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436(7054), 1161–1165. <http://dx.doi.org/10.1038/nature03867>. nature03867 [pii].
- 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.
- Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European Journal of Neuroscience*, 33(3), 530–538. <http://dx.doi.org/10.1111/j.1460-9568.2010.07527.x>.
- 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.
- 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.
- 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. <http://dx.doi.org/10.1152/jn.01114.2004>, 01114.2004 [pii].
- 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, N.Y.: 1991)*, 16(6), 835–848. <http://dx.doi.org/10.1093/cercor/bhj027>. bhj027 [pii].
- Chandrasekaran, B., Gandour, J. T., & Krishnan, A. (2007). Neuroplasticity in the processing of pitch dimensions: a multidimensional scaling analysis of the mismatch negativity. *Restorative Neurology and Neuroscience*, 25(3–4), 195–210.
- 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. <http://dx.doi.org/10.1016/j.neuron.2006.06.030>.
- Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, 96(2), 341–370.
- Cowan, N. (1987). Auditory sensory storage in relation to the growth of sensation and acoustic information extraction. *Journal of Experimental Psychology: Human Perception and Performance*, 13(2), 204–215.
- 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. <http://dx.doi.org/10.1523/JNEUROSCI.0470-08.2008>, 28/25/6430 [pii].
- Denham, S. (2005). Pitch detection of dynamic iterated rippled noise by humans and a modified auditory model. *Biosystems*, 79(1–3), 199–206.
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16(5), 419–423.
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36(2), 268–294. <http://dx.doi.org/10.1016/j.wocn.2007.06.005>.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: a dynamic dual pathway model. *Brain and Language*, 89(2), 267–276. [http://dx.doi.org/10.1016/S0093-934X\(03\)00351-1](http://dx.doi.org/10.1016/S0093-934X(03)00351-1). S0093934X03003511 [pii].
- 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. <http://dx.doi.org/10.1038/nn1141>.
- Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149–175.

- Gandour, J. T. (1994). Phonetics of tone. In R. Asher, & J. Simpson (Eds.), *The encyclopedia of language & linguistics* (Vol. 6, pp. 3116–3123). New York: Pergamon Press.
- Gandour, J. T. (2006). Brain mapping of Chinese speech prosody. In P. Li, L. H. Tan, E. Bates, & O. J. L. Tzeng (Eds.), *Handbook of East Asian psycholinguistics: Vol. 1. Chinese* (pp. 308–319). Cambridge, UK: Cambridge University Press.
- Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Language and Speech*, 21(1), 1–33.
- Gandour, J. T., & Krishnan, A. (2014). Neural bases of lexical tone. In H. Winkler, & P. Padakannaya (Eds.), *Handbook of South and Southeast Asian psycholinguistics* (pp. 339–349). Cambridge, UK: Cambridge University Press.
- Gilbert, C. D., & Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, 54(5), 677–696. <http://dx.doi.org/10.1016/j.neuron.2007.05.019>.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.
- 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.
- 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. <http://dx.doi.org/10.1016/j.cub.2010.04.044>.
- 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.
- 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. <http://dx.doi.org/10.1006/nimg.2001.0949>.
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *Neuroimage*, 22(2), 755–766. <http://dx.doi.org/10.1016/j.neuroimage.2004.01.025>. S1053811904000680 [pii].
- Hari, R., Pelizzone, M., Mäkelä, 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.
- Hertrich, I., Mathiak, K., Lutzenberger, W., & Ackermann, H. (2000). Differential impact of periodic and aperiodic speech-like acoustic signals on magnetic M50/M100 fields. *Neuroreport*, 11(18), 4017–4020.
- Huang, T., & Johnson, K. (2011). Language specificity in speech perception: perception of Mandarin tones by native and nonnative listeners. *Phonetica*, 67, 243–267.
- 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. <http://dx.doi.org/10.1111/j.1460-9568.2012.08278.x>.
- Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, 123(Pt 1), 155–163.
- Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in Neurobiology*, 82(3), 109–121. <http://dx.doi.org/10.1016/j.pneurobio.2007.03.005>. S0301-0082(07)00073-1 [pii].
- Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, 35(1), 104–117. <http://dx.doi.org/10.1016/j.jocp.2005.10.003>.
- Klatt, D. (1973). Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *Journal of the Acoustical Society of America*, 53(1), 8–16.
- 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. <http://dx.doi.org/10.1016/j.brainresrev.2007.07.021>. S0165-0173(07)00187-7 [pii].
- Kratochvil, P. (1985). Variable norms of tones in Beijing prosody. *Cahiers de Linguistique Asie Orientale*, 14(2), 153–174.
- 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. <http://dx.doi.org/10.1016/j.heares.2010.04.016>.
- 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. <http://dx.doi.org/10.1016/j.neuropsychologia.2012.08.013>.
- 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. <http://dx.doi.org/10.1016/j.bandl.2009.03.005>. S0093-934X(09)00042-X [pii].
- Krishnan, A., Gandour, J. T., Ananthakrishnan, S., & Vijayaraghavan, V. (2014). Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*, 59, 1–12. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.04.006>.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, 23(1), 81–95. <http://dx.doi.org/10.1016/j.jneuroling.2009.09.001>.
- 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. <http://dx.doi.org/10.1097/WNR.0b013e328353764d>.
- 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. <http://dx.doi.org/10.1016/j.bandl.2010.05.004>.
- 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. <http://dx.doi.org/10.1162/jocn.2009.21077>.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research. Cognitive Brain Research*, 25(1), 161–168. <http://dx.doi.org/10.1016/j.cogbrainres.2005.05.004>. S0926-6410(05)00123-0 [pii].
- 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. <http://dx.doi.org/10.1016/j.bandl.2013.11.006>.

- 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.
- Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neurophysiology*, *60*(6), 1799–1822.
- Lee, C. C., & Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, *14*(1), 108–114. <http://dx.doi.org/10.1038/nn.2713>.
- 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. <http://dx.doi.org/10.1038/nn737>.
- 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. <http://dx.doi.org/10.1016/j.neuroimage.2005.10.034>.
- 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.
- 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.
- Maess, B., Jacobsen, T., Schroger, E., & Friederici, A. D. (2007). Localizing pre-attentive auditory memory-based comparison: magnetic mismatch negativity to pitch change. *NeuroImage*, *37*(2), 561–571.
- Meyer, M. (2008). Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitschrift für Neuropsychologie*, *19*(2), 101–115.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, *102*(3), 1864–1877.
- 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. <http://dx.doi.org/10.1073/pnas.0701498104>, 0701498104 [pii].
- 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. <http://dx.doi.org/10.1371/journal.pbio.0060126>, 07-PLBI-RA-2244 [pii].
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Oxenham, A. J. (2012). Pitch perception. *Journal of Neuroscience*, *32*(39), 13335–13338. <http://dx.doi.org/10.1523/JNEUROSCI.3815-12.2012>.
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*(2), 97–107. <http://dx.doi.org/10.1038/nrn1603>, nrn1603 [pii].
- 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.
- 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. <http://dx.doi.org/10.1523/JNEUROSCI.0383-04.2004>, 24/30/6810 [pii].
- Pike, K. L. (1948). *Tone languages*. Ann Arbor, MI: University of Michigan Press.
- Plack, C. J., Carlyon, R. P., & Viemeister, N. F. (1995). Intensity discrimination under forward and backward masking: role of referential coding. *Journal of the Acoustical Society of America*, *97*(2), 1141–1149.
- Plack, C. J., Oxenham, A. J., & Fay, R. R. (Eds.). (2005). *Pitch: Neural coding and perception* (Vol. 24). New York: Springer.
- Plack, C. J., Turgeon, M., Lancaster, S., Carlyon, R. P., & Gockel, H. E. (2011). Frequency discrimination duration effects for Huggins pitch and narrowband noise (L). *Journal of the Acoustical Society of America*, *129*(1), 1–4. <http://dx.doi.org/10.1121/1.3518745>.
- Plack, C. J., & White, L. J. (2000). Perceived continuity and pitch perception. *Journal of the Acoustical Society of America*, *108*(3 Pt 1), 1162–1169.
- 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.
- 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), 1071–1086. <http://dx.doi.org/10.1098/rstb.2007.2160>, TM425571U1117682 [pii].
- 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.
- Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, *110*(2), 81–94. <http://dx.doi.org/10.1016/j.bandl.2008.12.001>.
- Ritter, S., Gunter Dosch, H., Specht, H. J., & Rupp, A. (2005). Neuromagnetic responses reflect the temporal pitch change of regular interval sounds. *Neuroimage*, *27*(3), 533–543. <http://dx.doi.org/10.1016/j.neuroimage.2005.05.003>.
- 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.
- Sayles, M., & Winter, I. M. (2007). The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Research*, *1171*, 52–66. <http://dx.doi.org/10.1016/j.brainres.2007.06.098>, S0006-8993(07)01509-0 [pii].
- 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. <http://dx.doi.org/10.1007/s00221-008-1286-z>.
- 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. <http://dx.doi.org/10.1016/j.heares.2006.01.003>.
- Soeta, Y., & Nakagawa, S. (2008). The effects of pitch and pitch strength on an auditory-evoked N1m. *Neuroreport*, *19*(7), 783–787. <http://dx.doi.org/10.1097/WNR.0b013e3282fe2085>.
- Soeta, Y., Nakagawa, S., & Matsuoka, K. (2005). Effects of the critical band on auditory-evoked magnetic fields. *Neuroreport*, *16*(16), 1787–1790. <http://dx.doi.org/10.1097/01.Wnr.0000185961.88593.4f>.
- 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*(10), 1892–1902.

- 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.
- Swaminathan, J., Krishnan, A., Gandour, J. T., & Xu, Y. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Transactions on Biomedical Engineering*, *55*(1), 281–287. <http://dx.doi.org/10.1109/TBME.2007.896592>.
- Tsang, Y. K., Jia, S., Huang, J., & Chen, H. C. (2011). ERP correlates of pre-attentive processing of Cantonese lexical tones: the effects of pitch contour and pitch height. *Neuroscience Letters*, *487*(3), 268–272. <http://dx.doi.org/10.1016/j.neulet.2010.10.035>. S0304-3940(10)01374-1 [pii].
- 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. <http://dx.doi.org/10.1016/j.heares.2010.04.015>.
- Wang, X. D., Wang, M., & Chen, L. (2013). Hemispheric lateralization for early auditory processing of lexical tones: dependence on pitch level and pitch contour. *Neuropsychologia*, *51*(11), 2238–2244. <http://dx.doi.org/10.1016/j.neuropsychologia.2013.07.015>.
- 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.
- 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.
- White, L. J., & Plack, C. J. (1998). Temporal processing of the pitch of complex tones. *Journal of the Acoustical Society of America*, *103*(4), 2051–2063.
- White, L. J., & Plack, C. J. (2003). Factors affecting the duration effect in pitch perception for unresolved complex tones. *Journal of the Acoustical Society of America*, *114*(6 Pt 1), 3309–3316.
- Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends in Neurosciences*, *28*(5), 255–263. <http://dx.doi.org/10.1016/j.tins.2005.03.009>.
- Winter, I. M., Wiegrebe, L., & Patterson, R. D. (2001). The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *Journal of Physiology*, *537*(Pt 2), 553–566.
- Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(4), 565–585.
- 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*(4), 420–422. <http://dx.doi.org/10.1038/nn1872>. nn1872 [pii].
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223–231. <http://dx.doi.org/10.1016/j.neuroscience.2010.06.077>. S0306-4522(10)00949-8 [pii].
- 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. <http://dx.doi.org/10.1016/j.neubiorev.2008.10.006>.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*, 61–83.
- Xu, Y. (2001). Pitch targets and their realization: evidence from Mandarin Chinese. *Speech Communication*, *33*, 319–337.
- Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, *120*(2), 1063–1074.
- Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, *111*(3), 1399–1413.
- Yip, M. (2002). *Tone*. New York: Cambridge University Press.
- Yrttiaho, S., Alku, P., May, P. J., & Tiitinen, H. (2009). Representation of the vocal roughness of aperiodic speech sounds in the auditory cortex. *Journal of the Acoustical Society of America*, *125*(5), 3177–3185. <http://dx.doi.org/10.1121/1.3097471>.
- Yrttiaho, S., Tiitinen, H., Alku, P., Miettinen, I., & May, P. J. (2010). Temporal integration of vowel periodicity in the auditory cortex. *Journal of the Acoustical Society of America*, *128*(1), 224–234. <http://dx.doi.org/10.1121/1.3397622>.
- 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. <http://dx.doi.org/10.1121/1.2888489>.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, *84*(2), 566–572.
- Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS Biology*, *10*(7), e1001372. <http://dx.doi.org/10.1371/journal.pbio.1001372>.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*(10), 946–953.
- 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. <http://dx.doi.org/10.1098/rstb.2007.2161>. J412P80575385013 [pii].
- Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS One*, *6*(6), e20963. <http://dx.doi.org/10.1371/journal.pone.0020963>. <http://www.ncbi.nlm.nih.gov/pubmed/21695133>.