Brain Research

CrossMark

Research report

Language-experience plasticity in neural representation of changes in pitch salience

Ananthanarayan Krishnan, Jackson T. Gandour*, Chandan H. Suresh

Department of Speech Language Hearing Sciences, Purdue University, Lyles Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907-2122, USA

ARTICLE INFO

ABSTRACT

Neural representation of pitch-relevant information at the brainstem and cortical levels of processing is influenced by language experience. A well-known attribute of pitch is its *salience*. Brainstem frequency following responses and cortical pitch specific responses, recorded concurrently, were elicited by a pitch salience continuum spanning weak to strong pitch of a dynamic, iterated rippled noise pitch contour—homolog of a Mandarin tone. Our aims were to assess how language experience (Chinese, English) affects i) enhancement of neural activity associated with pitch salience at brainstem and cortical levels, ii) the presence of asymmetry in cortical pitch representation, and iii) patterns of relative changes in magnitude along the pitch salience continuum. Peak latency (Fz: Na, Pb, and Nb) was shorter in the Chinese than the English group across the continuum. Peak-to-peak amplitude (Fz: Na–Pb, Pb–Nb) of the Chinese group grew larger with increasing pitch salience, but an experience-dependent advantage was limited to the Na–Pb component. At temporal sites (T7/T8), the larger amplitude of the Chinese group across the continuum was both limited to the Na–Pb component and the right temporal site. At the brainstem level, F0 magnitude gets larger as you increase pitch salience, and it too reveals Chinese superiority. A direct comparison of cortical and brainstem responses for the Chinese group reveals different patterns of relative changes in magnitude along the pitch salience continuum. Such differences may point to a transformation in pitch processing at the cortical level presumably mediated by local sensory and/or extrasensory influence overlaid on the brainstem output.

1. Introduction

Pitch is a perceptual attribute that plays an important role in the perception of speech, language and music. For many types of complex sounds, including speech and music, pitch and its *salience* is closely related to the temporal periodicity strength in the stimulus waveform fine structure (Shofner, 2002; Yost, 1996a). Iterated rippled noise (IRN) is a complex sound that permits systematic manipulation of the temporal fine structure, and therefore the magnitude of pitch salience. IRN is produced by

adding a delayed copy of random noise to the original noise and then repeating this delay-and-add process *n* times (Yost, 1996b). Increasing the number of iterations produces an increase in temporal regularity of the noise and a spectral ripple in its long-term power spectrum. Perceptually, IRN yields a pitch corresponding to the reciprocal of the delay (Patterson et al., 1996). Its corresponding pitch salience grows with increasing number of iterations. The increase in pitch salience with increasing temporal regularity of the IRN stimulus is correlated with an increase in pitch-relevant neural activity in both cortical and subcortical auditory neurons as evidenced by data from physiological (Griffiths et al., 1998; Sayles and Winter, 2007); electrophysiological (Krishnan et al., 2010a; Krumbholz et al., 2003; Soeta et al., 2005); functional brain imaging (Griffiths et al., 2001); and intracranial electrode recordings (Schonwiesner and Zatorre, 2008).

Pitch provides an excellent window for studying experience-dependent effects on both brainstem and cortical components of a well-coordinated, hierarchical processing network. There is growing empirical evidence to support the notion that the neural representation of pitch relevant information at both brainstem and cortical levels of processing is influenced by one's long-term experience or short-term training with language and music (see Chandrasekaran and Kraus, 2010; Gandour and Krishnan, 2014; Krishnan and Gandour, 2014; Patel, 2008; Zatorre and Baum, 2012, for reviews). There is neuroanatomical evidence of ascending and descending pathways (Saldana et al., 1996) and physiological evidence of improved signal representation in subcortical structures mediated by the corticofugal system (Suga et al., 2003). These findings suggest that neural processes mediating experience-dependent plasticity for pitch at the brainstem and cortical levels may be well-coordinated. The effects of musical training, for example, yield a correlation between brainstem and cortical responses (Musacchia et al., 2008), implying that neural representations of pitch, timing and timbre cues at the two levels are shaped in a coordinated manner through corticofugal modulation of subcortical afferent circuitry. However, little is known about how language experience shapes specific attributes of pitch at each level of the processing hierarchy or how it modulates the nature of the interplay between them. The scalp-recorded brainstem frequency following response (FFR) and the cortical pitch response (CPR) represent neural activity relevant to pitch at brainstem and cortical levels, respectively. As such, they provide physiologic windows to evaluate the hierarchical organization of pitch processing along the auditory pathway.

The short latency (7–12 ms) FFR reflects sustained phase-locked neural activity in a population of neural elements primarily within the rostral brainstem (Chandrasekaran and Kraus, 2010; Krishnan, 2007) with appreciable cortical contribution of phase-locked FFR-like activity with longer latencies limited to frequencies below 100 Hz (Herdman et al., 2002; Steinschneider et al., 1999). Pitch-relevant information preserved in the FFR is strongly correlated with perceptual pitch measures (Bidelman and Krishnan, 2011; Krishnan and Plack, 2009; Krishnan et al., 2010a; Parbery-Clark et al., 2009), suggesting that acoustic features relevant to pitch are already emerging in representations at the level of the brainstem. Of special relevance to this study, FFR neural periodicity strength increases with increasing pitch salience and accurately predicts the perceptual salience of IRN pitch (Krishnan et al., 2010a). FFRs further reveal that experience-dependent

brainstem mechanisms for pitch are especially sensitive to those attributes of pitch contours that provide cues of high perceptual saliency in degraded as well as normal listening conditions (Krishnan et al., 2010b).

The Na component of the CPR, an EEG correlate of the MEG-derived pitch onset response (POR), reflects pitch-specific synchronized neural activity in the auditory cortex. Source analysis of MEG derived pitch onset response (Gutschalk et al., 2002; Krumbholz et al., 2003)—corroborated by human depth electrode recordings (Griffiths et al., 2010; Schonwiesner and Zatorre, 2008) and source analysis of the EEG derived Na (Bidelman and Grall, 2014) indicates that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing (Johnsrude et al., 2000; Penagos et al., 2004). The CPR, in addition, is characterized by multiple transient components (Na: 125–150 ms, Pb: 200–220, Nb: 270–285 ms) that may index different temporal attributes of pitch contours (Krishnan et al., 2014a, 2014b). We have adopted the Krumbholz et al. experimental paradigm to EEG recordings in order to extract the CPR and FFR concurrently (Krishnan et al., 2012a). In response to IRN *steady-state* pitch stimuli, English monolinguals exhibit larger magnitude in response to strong as compared to weak pitch at both brainstem and cortical levels, i.e., more robust encoding for salient pitch. This change in response magnitude is strongly correlated with behavioral measures of change in perceptual pitch salience. As far as we know, no one has yet to carry out a systematic parametric evaluation of the nature of the interplay underlying processing of *dynamic* pitch stimuli between these two levels of pitch processing along the auditory pathway.

The aim of this study is to determine how systematic changes in pitch salience along a continuum going from weak to strong pitch of a dynamic, IRN high-rising pitch contour—a homolog of Mandarin Chinese Tone 2 (T2)—may alter the strength of the representation of pitch-relevant information preserved in the simultaneously recorded brainstem FFR and cortical CPR as a function of language experience (Chinese, English) (cf. Krishnan et al., 2012a, steady-state pitch). This gives us a unique window to examine the coordination between different stages of pitch processing in real time, which may otherwise be obscured by inferences drawn from separate evaluation of neural responses evoked by different stimulation/acquisition paradigms or comparisons across studies. We hypothesize that language experience enhances sensitivity to changes in pitch salience at the auditory brainstem and cortex, and preferentially recruits the right hemisphere for pitch processing.

A direct comparison of cortical and brainstem responses is expected to reveal different patterns of relative changes in magnitude along the pitch salience continuum. Such findings would implicate a transformation in the nature of processing between the two levels with respect to pitch salience.

2. Results

2.1. Cortical pitch response components

2.1.1. Response morphology

Fig. 1 (top) illustrates that Fz-linked(T7/T8)-derived CPR components (Na, Pb, and Nb) of the pitch-eliciting segment (color)
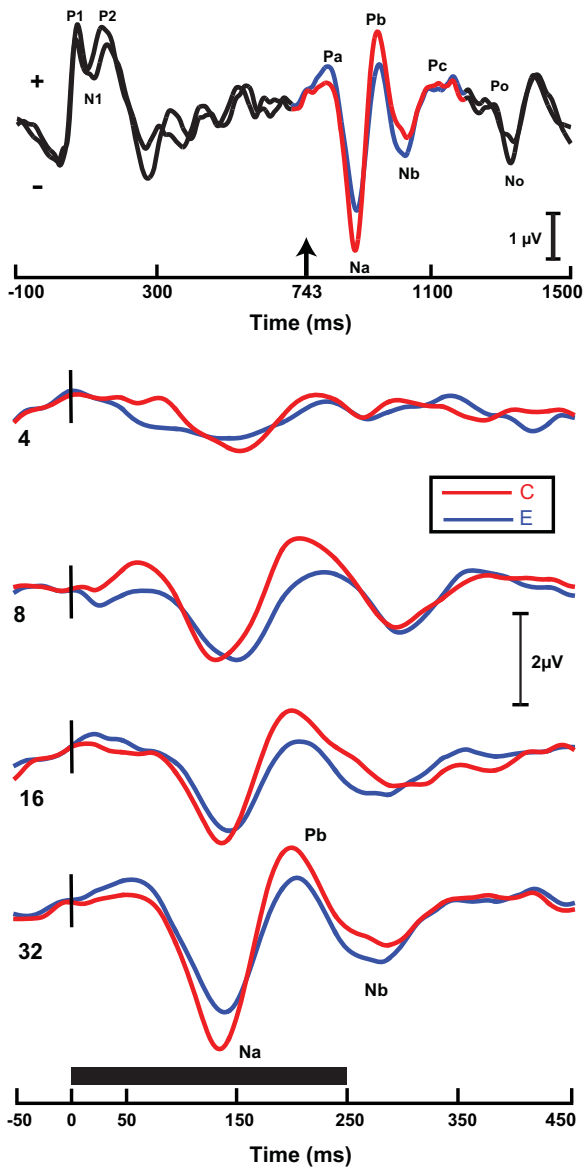
Fig. 1 – Grand average waveforms of the Chinese (C) and English (E) groups at the Fz electrode site per iteration step (4, 8, 16, 32). Waveforms consisting of three segments (top) illustrate the experimental paradigm used to acquire cortical responses: a 250 ms pitch segment ($n=32$) preceded and followed by 750 ms and 250 ms noise segments, respectively. No remarkable differences between groups are observed in their CPR responses during the noise segments (top: black portion). Na, Pb, and Nb (top: color portion) are the most robust pitch-relevant components. CPR waveforms elicited by the four stimuli (bottom) show that the amplitude of the pitch-relevant components (Na, Pb, and Nb) appear to be more robust for the Chinese as compared to the English especially in response to stimuli with higher iteration steps (8, 16, 32). 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.
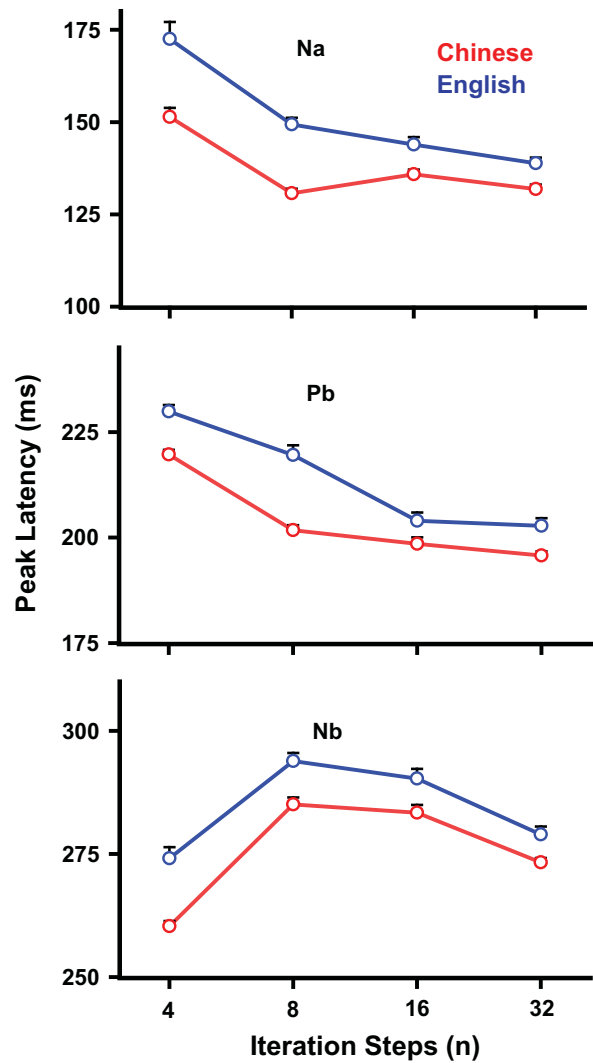

Fig. 2 – Mean peak latency CPR components (Na, Pb, and Nb) elicited by each of the four iteration steps at the Fz electrode site in both Chinese and English groups. In both Na (top) and Pb (middle), the stimulus with weak pitch salience (4) elicited a longer peak latency than the other three (8, 16, 32). In Nb (bottom), stimuli of medium pitch salience (8, 16) elicited longer peak latency as compared to either weak (4) or strong pitch salience (32) for both groups. In general, peak latencies of the English group are longer than those of Chinese regardless of CPR component or iteration step. No matter the language experience of the listener, peak latency gets shorter as the pitch becomes more salient, except for the weak iteration step (4) as reflected by Nb. CPR, cortical pitch response. Error bars = ±1 SE.

are clearly identifiable embedded within the stimulus paradigm. The pitch-eliciting segment is preceded and followed by noise segments (black). Fig. 1 (bottom) displays only the time window of the grand averaged CPR components per number of iteration steps. The amplitude of pitch-relevant components appears to increase with pitch salience and to be

more robust for the Chinese group at iteration steps 8, 16, and 32. Also, Na and Pb peak latencies appear to be shorter for the Chinese group across iterations.

2.1.2. Fz: latency (Na, Pb, Nb)

Fig. 2 displays mean peak latency of CPR components (Na, Pb, and Nb) elicited by the four stimuli across the pitch salience continuum. As reflected by Na (top), the ANOVA yielded a group × stimulus interaction ($F_{3,73}=9.88$, $p<0.0001$, $\eta_P^2=0.291$). In both groups, the stimulus with weak pitch salience (I4) elicited a longer response peak latency than the other three (C: I8 vs I4, $t_{73}=-10.59$, $p<0.0001$; I16 vs I4, $t=-8.09$; I32 vs I4, $t=-9.89$; E: I8 vs I4, $t=-10.06$; I16 vs I4, $t=-12.38$; I32 vs I4, $t=-14.54$). In addition, the stimulus with the next-to-weakest pitch salience (I8) evoked a longer latency as compared to the strongest (I32) in the English group only ($t=-5.00$, $p<0.0001$). Across stimuli, response peak latencies were longer in English than Chinese listeners (C vs E: I4, $t=-6.69$, $p<0.0001$; I8, $t=-6.96$; I16, $t=-2.87$, $p=0.0045$; I32, $t=-2.48$, and $p=0.0154$).

As reflected by Pb (middle), a similar group × stimulus interaction was observed ($F_{3,73}=9.68$, $p<0.0001$, $\eta_P^2=0.285$). Regardless of language affiliation, the stimulus with weak pitch salience (I4) elicited a longer response peak latency than the other three (C: I8 vs I4, $t_{73}=-9.94$, $p<0.0001$; I16 vs I4, $t=-11.88$; I32 vs I4, $t=-12.65$; E: I8 vs I4, $t=-5.19$; I16 vs I4, $t=-12.93$; I32 vs I4, $t=-13.53$). In addition, the stimulus with the next-to-weakest pitch salience (I8) evoked a longer latency as compared to the strongest, I32 (C: I32 vs I8, $t=-3.50$, $p=0.0422$; E: I32 vs I8, $t=-9.31$, $p<0.0001$). Also, I8 was longer than I16 for the English group only ($t=-8.64$, $p<0.0001$). Across stimuli, response peak latencies were longer in English than Chinese listeners (C vs E: I4, $t=-4.16$, $p<0.0001$; I8, $t=-7.86$; I16, $t=-2.76$, $p=0.0073$; I32, $t=-2.85$, $p=0.0057$).

As reflected by Nb (bottom), the ANOVA revealed main effects of group ($F_{1,26}=45.30$, $p<0.0001$, $\eta_P^2=0.344$) and stimulus ($F_{3,73}=76.10$, $p<0.0001$, $\eta_P^2=0.749$). Pooling across stimuli, response peak latencies were longer in English than Chinese listeners ($t_{26}=-6.73$, $p<0.0001$). Pooling across groups, stimuli of medium pitch salience (I8, I16) elicited longer peak latency as compared to the one with the strongest pitch salience, I32 (I32 vs I16, $t_{73}=-6.82$, $p<0.0001$; I32 vs I8, $t=-8.93$). In contrast, the stimulus with the strongest pitch salience, I32, evoked longer peak latency than I4 at the opposite end of the pitch salience continuum (I32 vs I4, $t=5.02$).

In general, these combined findings on Fz latency suggest that regardless of language affiliation, response peak latency gets shorter as the pitch becomes more salient. For Na and Pb, this pattern is evident across iteration steps for both language groups; for Nb, the pattern holds true for all but the weak iteration step (I4). Response peak latency is further modulated by language experience. Latencies are shorter in the Chinese as compared to the English group as reflected by all three CPR components (Na, Pb, and Nb).

2.1.3. Fz: amplitude (Na–Pb, Pb–Nb)

Fig. 3 displays mean peak-to-peak amplitude of CPR components (Na–Pb, Pb–Nb) across the pitch salience continuum. For Na–Pb (top), the ANOVA revealed main effects of group ($F_{1,26}=11.61$, $p=0.0021$, $\eta_P^2=0.186$) and stimulus ($F_{3,73}=78.18$, $p<0.0001$,
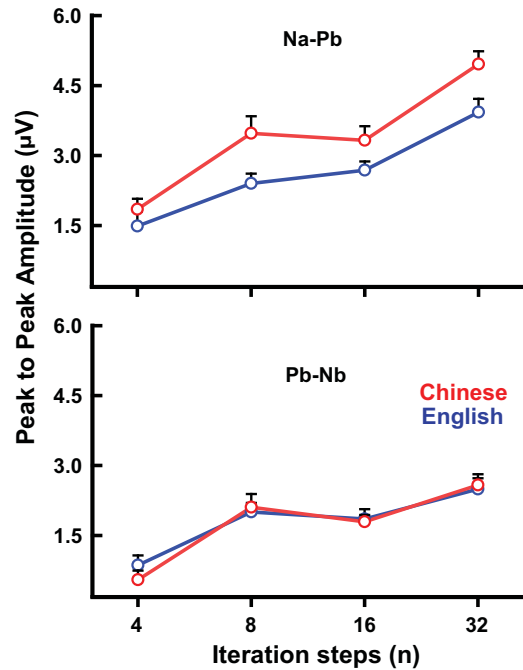

Fig. 3 – Mean peak-to-peak amplitude of CPR components (Na–Pb, Pb–Nb) elicited by each of the four iteration steps at the Fz electrode site in both Chinese and English groups. For Na–Pb (top), amplitude is larger for the Chinese group relative to the English across the pitch salience continuum; whereas no language group advantage is evident for Pb–Nb (bottom). The same pattern of increasing amplitude as a function of pitch salience, however, is observed in both groups across components. Stimuli with stronger pitch salience (8, 16, 32) have larger amplitude as compared to weak (4).

$\eta_P^2=0.766$). Irrespective of pitch salience, Chinese listeners exhibited larger amplitude relative to the English across all four stimuli (I4, I8, I16, I32). In general, both groups showed the same pattern of increasing amplitude as a function of stimulus pitch salience, including a leveling off for stimuli of medium pitch salience, I8 and I16 (I32 vs I16, $t_{73}=8.12$, $p<0.0001$; I32 vs I8, $t=8.12$; I32 vs I4, $t=8.83$; I16 vs I4, $t=7.58$; I8 vs I4, $t=6.92$). For Pb–Nb (bottom), the ANOVA yielded a main effect of stimulus only ($F_{3,73}=33.00$, $p<0.0001$, $\eta_P^2=0.580$). Post-hoc multiple comparisons indicated that the three stimuli of stronger pitch salience (I8, I16, I32) had larger amplitude than I4, the stimulus with weak pitch salience (I32 vs I4, $t=9.71$, $p<0.0001$; I16 vs I4, $t=5.72$; I8 vs I4, $t=7.20$). In addition, the stimulus with the strongest pitch salience (I32) also exhibited larger amplitude than those of medium pitch salience, I16 and I8 (I32 vs I16, $t=4.25$, $p<0.0001$; I32 vs I8, $t=2.68$, $p=0.0550$, marginal). These combined findings on Fz amplitude reveal that the effects of language experience on sensitivity to changes in pitch salience are restricted to Na–Pb.

2.1.4. T7/T8 and F3F4: amplitude of CPR components

Grand average waveforms of the CPR components (left) in response to each of the four iteration steps along the pitch salience continuum per language group and their corresponding spectra (right) are displayed in Fig. 4. A rightward
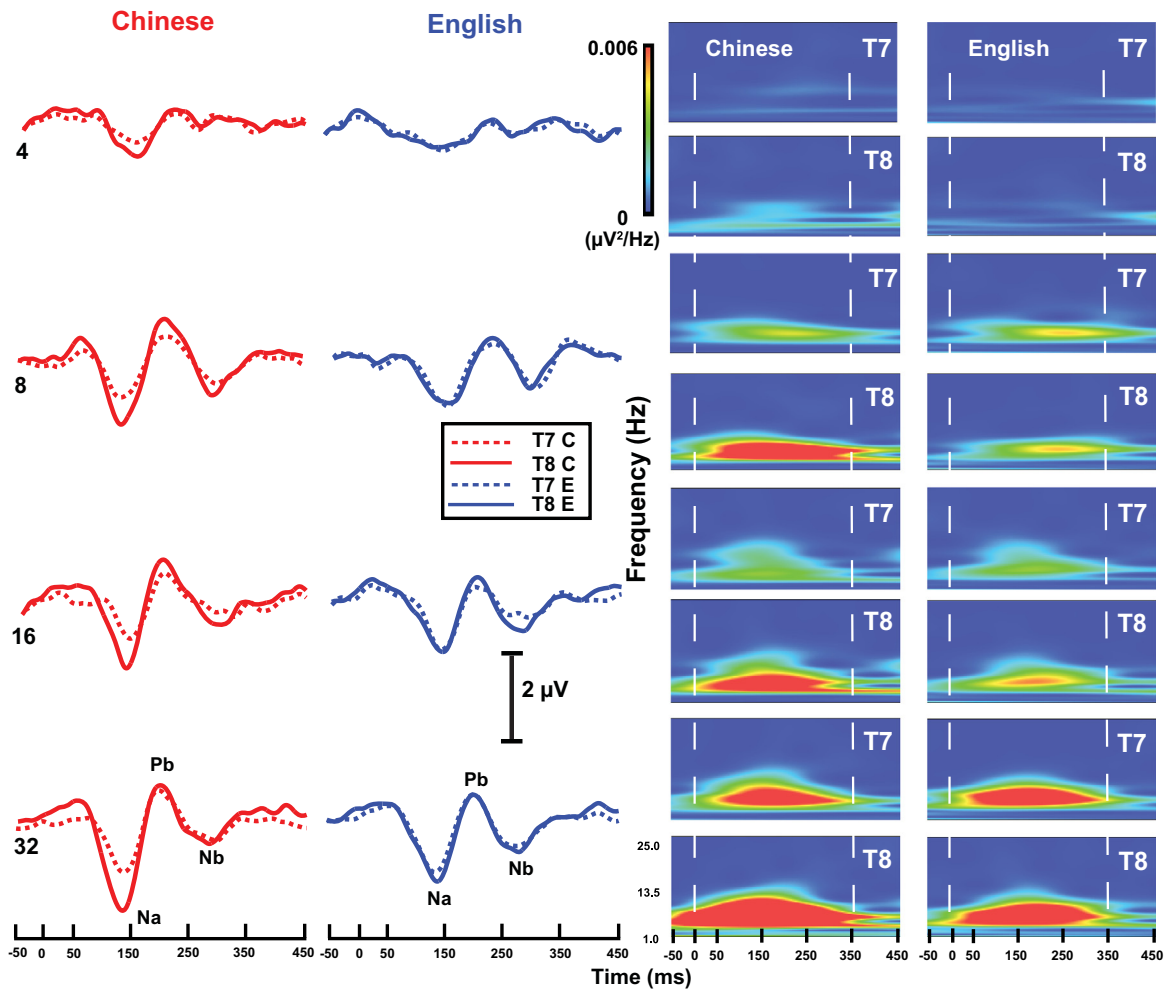
Fig. 4 – Grand average waveforms (left) and their corresponding spectra (right) of the CPR components for the two language groups (Chinese, red; English, blue) recorded at electrode sites T7 (dashed) and T8 (solid) for each of the four stimuli (4, 8, 16, 32). CPR waveforms appear to show a rightward asymmetry in favor of the right temporal electrode site for the Chinese group across iteration steps (4, 8, 16, 32). The robust rightward asymmetry is clearly evident in the spectrotemporal plots throughout the pitch salience continuum. Both Na–Pb and Pb–Nb time windows are included between the two, vertical white dashed lines (right). The zero on the *x*-axis denotes the time of onset of the pitch-eliciting segment of the four stimuli.

asymmetry (T8>T7) is apparent in CPR components of the Chinese group as evidenced in both waveforms and spectrotemporal plots across iteration steps. In contrast to T7/T8, frontal electrode sites (F3/F4; Supplementary Fig. S1) reveal no discernible difference between language groups either in their grand-averaged waveforms or spectrotemporal plots.

T7/T8 peak-to-peak amplitude of Na–Pb (top) and Pb–Nb (bottom) per iteration step are displayed for each language group and electrode site in Fig. 5 (F3/F4; Supplementary, Fig. S2). As reflected by Na–Pb, the three-way (group × stimulus × electrode site) ANOVA revealed a main effect of stimulus ($F_{3,66}=33.64$, $p<0.0001$, $\eta_P^2 = 0.363$) and a group × electrode site interaction ($F_{1,91}=17.81$, $p<0.0001$, $\eta_P^2 = 0.142$). Regardless of group, T7/T8 amplitude increased in response to increasing iterations, including a plateau for stimuli with medium pitch salience, I8 and I16 (I32 vs I16, $t_{66}=6.16$, $p<0.0001$; I32 vs I8, $t=6.23$; I32 vs I4, $t=9.38$; I16 vs I4, $t=4.96$; I8 vs I4, $t=4.84$). The interaction between group and electrode site revealed

language-dependent effects. A rightward asymmetry was found in the Chinese group (T7 vs T8: $t_{91}=-7.07$, $p<0.0001$), but not the English. Moreover, Chinese listeners' response amplitudes were larger than those of English over the right electrode site only (C vs E: $t=3.33$, $p=0.0013$).

As reflected by Pb–Nb, the ANOVA revealed main effects of stimulus ($F_{3,66}=16.02$, $p<0.0001$, and $\eta_P^2 = 0.270$) and electrode site ($F_{1,91}=14.40$, $p=0.0003$, and $\eta_P^2 = 0.134$). Similar to Na–Pb, T7/T8 amplitude increased in response to increasing iterations, including a plateau for the middle two iteration steps, I8 and I16 (I32 vs I4, $t_{66}=6.59$, $p<0.0001$; I16 vs I4, $t=3.65$, $p=0.0013$; I8 vs I4, $t=5.08$, $p<0.0001$; I32 vs I16, $t=4.04$, $p=0.0009$). The electrode site main effect revealed a preference for the right temporal electrode site irrespective of stimulus or group (T7 vs T8: $t_{91}=-3.79$, $p=0.0003$). Taken together, these findings on T7/T8 amplitude indicate that the effects of language experience on sensitivity to changes in pitch salience are not only restricted to Na–Pb, but also are limited to the right temporal electrode site.
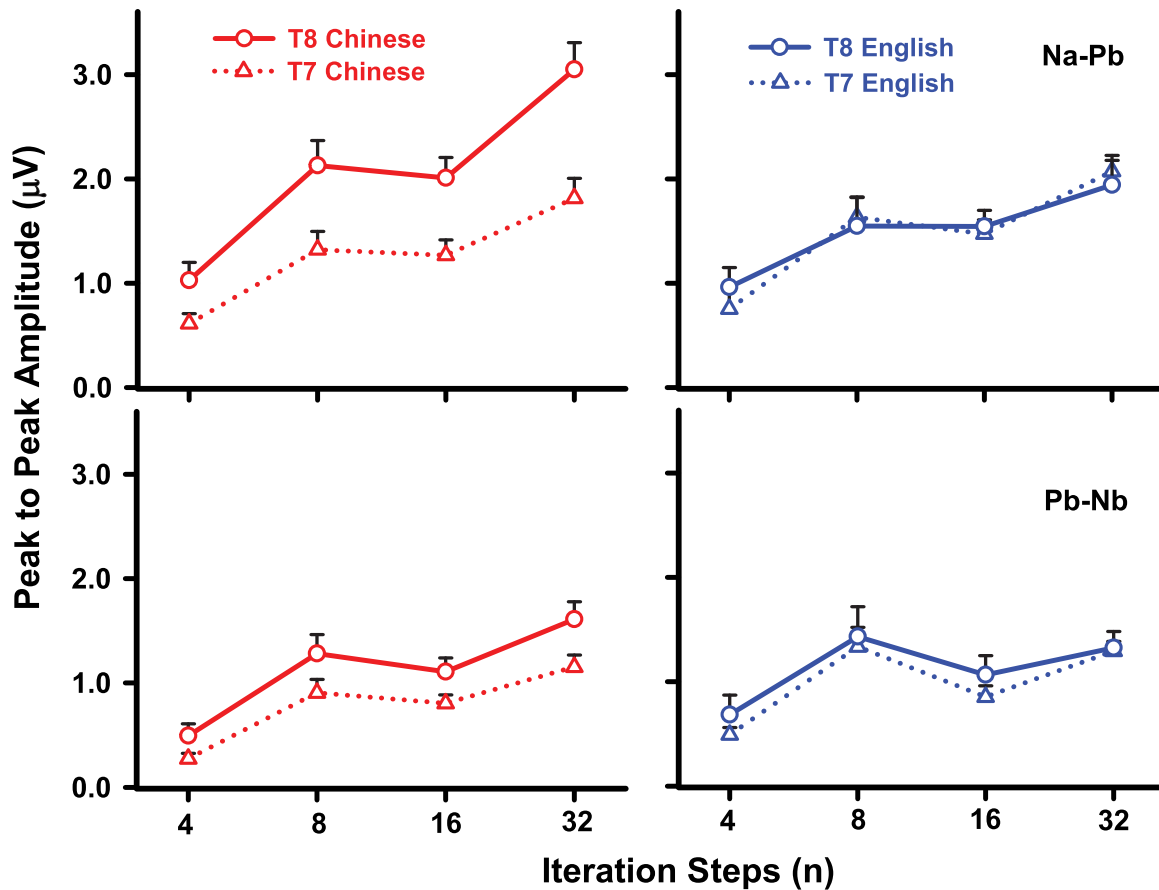
Fig. 5 – Mean peak-to-peak amplitude of CPR components (Na–Pb, Pb–Nb) elicited by each of the four iteration steps at the T7/T8 temporal electrode sites in both Chinese and English groups. For both Na–Pb (top row) and Pb–Nb (bottom row), amplitude steadily increases from weak (4) to strong (32) pitch salience for both language groups, including a plateau for medium pitch salience (8, 16). In the case of Na–Pb, however, a right-sided asymmetry is found in the Chinese group only. A language-experience effect manifests itself at the right temporal electrode site; Chinese amplitudes are larger than those of English. In contrast, a rightward asymmetry is reflected by Pb–Nb irrespective of iteration step or group. Thus, the effects of language experience on sensitivity to changes in pitch salience are limited to the Na–Pb component at the right temporal electrode site.

2.2. Brainstem pitch response component

2.2.1. Temporal and spectral response characteristics

Grand averaged FFR waveforms (left panel) and their spectra (right panel) are shown as a function of iteration steps for Chinese and English listeners in Fig. 6. Qualitatively, FFR waveforms show clearer periodicity and larger amplitude with increasing iterations (4, 8, 16, and 32). FFR spectra, particularly for the Chinese, likewise, reveal clearer and more robust spectral components at the F0 with increasing pitch salience. Overall, the responses appear to be more robust for the Chinese compared to the English.

2.2.2. Neural pitch strength as a function of iteration step

FFR encoding of F0 at each of the four iteration steps along the pitch salience continuum are shown in Fig. 7. An omnibus ANOVA revealed significant main effects of group ($F_{1,26} = 6.83$, $p = 0.0147$, $\eta_P^2 = 0.857$) and stimulus ($F_{3,78} = 29.13$, $p < 0.0001$, $\eta_P^2 = 0.967$). Chinese had larger F0 magnitudes than English across all four iteration steps, indicating an experience-dependent enhancement of pitch encoding in the brainstem regardless of the degree of pitch salience. Pooling across groups, post-hoc

multiple comparisons of stimuli showed that the stimulus with the strongest pitch salience (I32) had larger F0 magnitude than the three other stimuli with lesser degrees of pitch salience (I32 vs I16, $t_{78} = 4.85$, $p < 0.0001$; I32 vs I8, $t = 8.37$; I32 vs I4, $t = 7.75$). The stimulus with the next-to-strongest pitch salience (I16), in turn, evoked larger F0 magnitude than the two stimuli (I8, I4) with weaker pitch salience (I16 vs I8, $t = 4.85$, $p = 0.0043$; I16 vs I4, $t = 2.90$, $p = 0.0290$). These findings suggest that brainstem neural pitch strength increases systematically with increasing temporal regularity in stimulus periodicity, indicating more robust encoding for salient pitch. Chinese superiority notwithstanding, the absence of an interaction between group and stimulus indicates that pitch salience is encoded by a brainstem pitch mechanism shared in common across languages.

2.3. Comparison of cortical and brainstem responses

Fig. 8 shows normalized magnitude ratios of cortical (Na–Pb; top; Pb–Nb, middle) and brainstem (FFR; bottom) responses derived from successive iteration steps (I4–I8; I8–I16; I16–I32) in both Chinese and English groups. Regardless of group, we
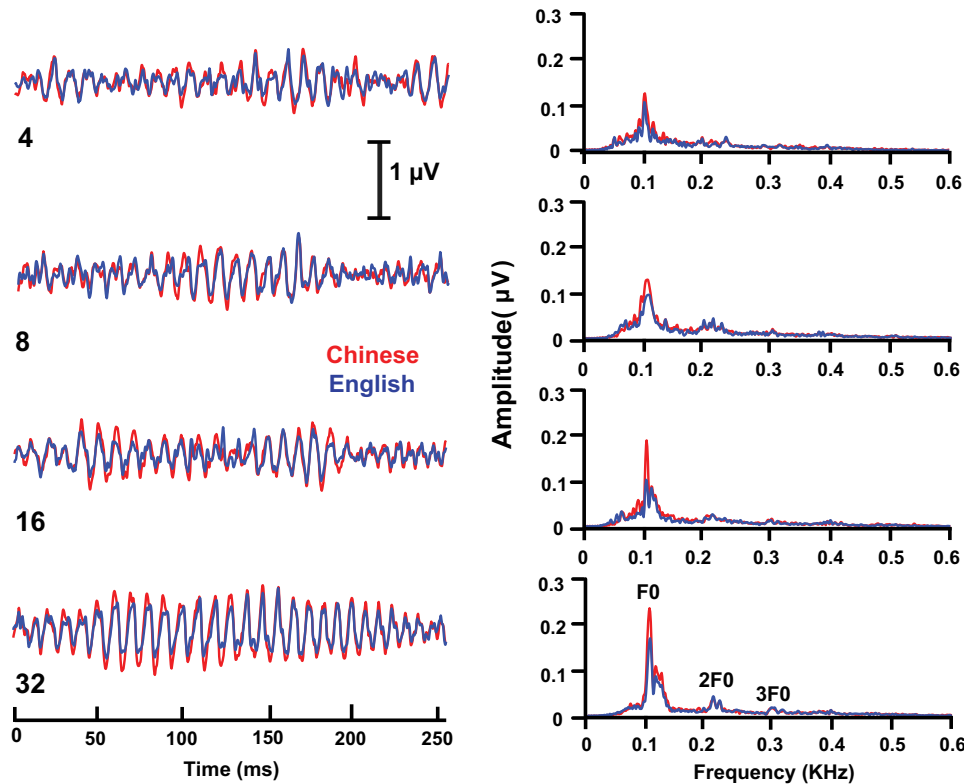
Fig. 6 – FFR waveforms (left) and spectra (right) as a function of iteration steps (I4, I8, I16, I32) computed from grand averaged brainstem responses for Chinese (red) and English (blue) listeners. Waveform periodicity is markedly improved at higher iteration steps. Response spectra show robust phase-locked energy at the fundamental frequency (F0) and its integer- related harmonics (2F0, 3F0). FFR, fundamental frequency following response. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
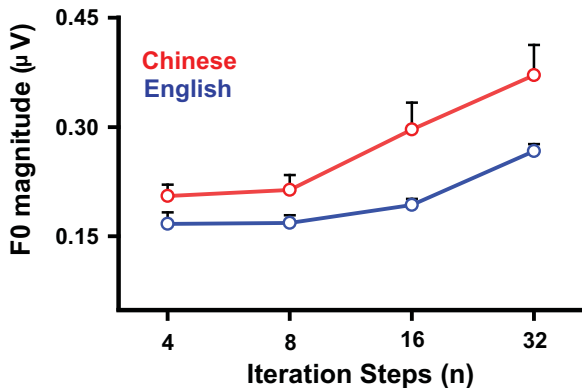

Fig. 7 – Mean F0 magnitude extracted from the brainstem FFR elicited by each of the four iteration steps in both Chinese and English groups. Chinese had larger F0 magnitudes than English across the pitch salience continuum. The stimulus with strong pitch salience (32) has larger F0 magnitude than the other three (4, 8, 16). Regardless of language experience, neural pitch strength increases systematically with increasing temporal regularity in stimulus periodicity, thus providing a robust measure of pitch salience at the level of the brainstem.

observe a monotonic relationship between successive iteration steps at the brainstem level, meaning that the representation of pitch gets stronger with increasing salience of pitch. The relationship between successive iteration steps at the

cortical level, however, is not monotonic. Omnibus ANOVAs yielded interactions between component and stimulus in both Na–Pb and FFR ($F_{2,35}=11.20$, $p=0.0002$, $\eta_P^2=0.341$) and Pb–Nb and FFR ($F_{2,35}=21.46$, $p<0.0001$, $\eta_P^2=0.552$). No main or interaction effect involving group reached significance, meaning that the observed effects apply irrespective of language experience.

At the simple effect level of cortical component, the size of the change in magnitude ratio of pitch salience was greater from weak to medium-weak (I4–I8) and from medium-strong to strong (I16–I32) when compared to medium-weak vs medium-strong (I8–I16) (Na–Pb: I4–18 vs I8-I16, $t_{35}=4.18$, $p=0.0005$; I8–I16 vs I16–I32, $t_{35}=-3.25$, $p=0.0077$; Pb–Nb: I4–18 vs I8–I16, $t_{35}=6.03$, $p<0.0001$; I8–I16 vs I16–I32, $t_{35}=-3.03$, $p=0.0137$). As reflected by Pb–Nb, the change in magnitude ratio of pitch salience from weak to medium-weak was also greater than from medium-strong to strong (I4–I8 vs I16–I32, $t_{35}=3.15$, $p=0.0099$). At the simple effect level of the brainstem component (FFR), just the opposite was the case. The size of the change in magnitude ratio of pitch salience was greater from medium-strong to strong (I16–I32) than from weak to medium-weak (I4–I8) (Na–Pb: I4–I8 vs I16–I32, $t_{35}=-3.81$, $p=0.0016$; Pb–Nb: I4–I8 vs I16–I32, $t_{35}=-3.41$, $p=0.0050$). At the simple effect level of stimulus, the change from weak to medium-weak pitch salience (I4–I8) elicited a stronger magnitude ratio in the cortical components relative to the brainstem (Na–Pb vs FFR: $t_{35}=4.97$, $p<0.0001$; Pb–Nb vs
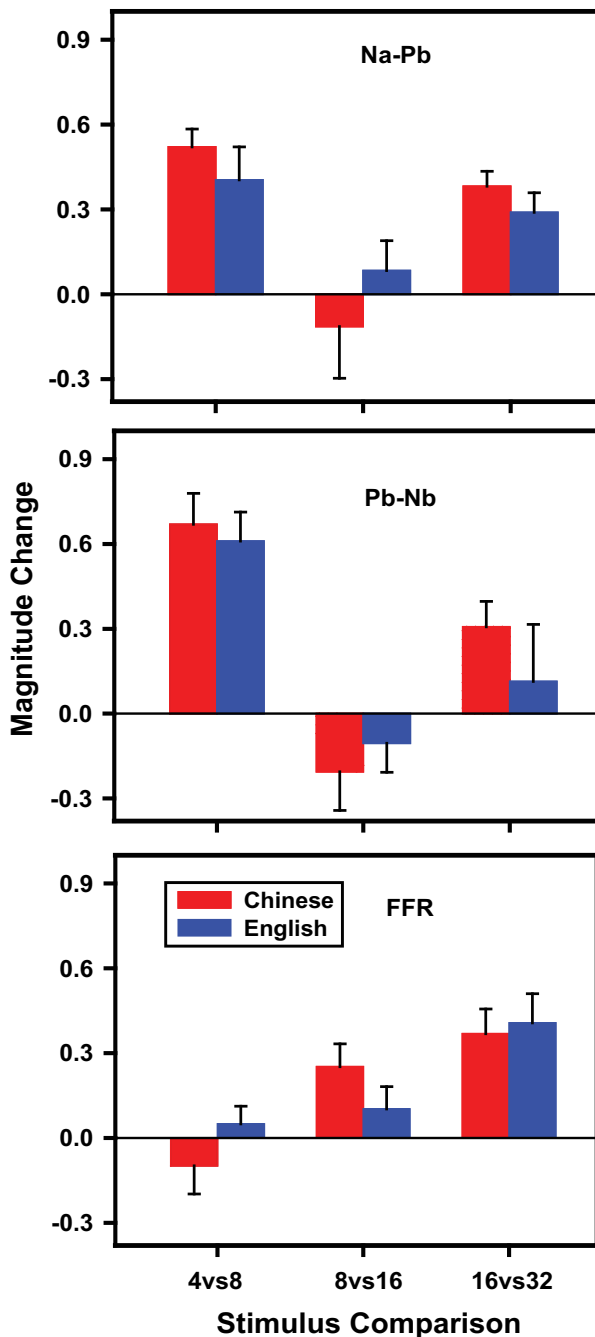
Fig. 8 – Mean normalized magnitude ratio of cortical (Na–Pb, top; Pb–Nb, middle) versus brainstem (FFR, bottom) responses derived from successive iteration steps (I4–I8; I8–I16; I16–I32) in both Chinese and English groups. Regardless of group, the pattern of differences in magnitude ratio elicited by successive iteration steps is virtually identical in both Na–Pb and FFR and Pb–Nb and FFR. The brainstem pattern is monotonic; the cortical, non-monotonic. At the cortical level, either weak to medium-weak (4 vs 8) pitch salience or medium-strong to strong (16 vs 32) is larger than that of medium-weak to medium-strong (8 vs16). At the brainstem level, strong pitch salience (16 vs 32) is greater than that involving weak (4 vs 8). For weaker pitch salience (4 vs 8), cortical components have larger magnitude ratios than the brainstem; conversely, the brainstem has larger magnitude ratios for intermediate steps (8 vs16).

FFR: $t_{35}=6.16$, $p<0.0001$). In contrast, the change from medium-weak to medium-strong pitch salience (I8–I16) elicited a stronger magnitude ratio in the brainstem than in the cortex (Na–Pb vs FFR: $t_{35}=-2.11$, $p<0.0418$; Pb–Nb vs FFR: $t_{35}=-3.49$, $p=0.0013$). These findings taken together suggest that changes in magnitude between successive iteration steps along the pitch salience continuum are not the same for cortical and brainstem responses. Such differences in responses are likely to implicate differences in the nature of processing between the two levels with respect to pitch salience.

3. Discussion

The major findings of this study demonstrate that pitch-related neural activity as reflected in scalp-recorded cortical and brainstem responses varies as a function of pitch salience and language experience. Fz peak latency (Na, Pb, and Nb) gets shorter as you increase pitch salience, but is modulated by language experience (C<E). Fz amplitude (Na–Pb, Pb–Nb) similarly gets larger as you increase pitch salience, but experience-dependent sensitivity to changes in pitch salience is limited to Na–Pb (C>E). At temporal sites, experience-dependent sensitivity to changes in pitch salience are also limited to Na–Pb (C>E) and, in addition, to the right temporal site (T8>T7). At the brainstem level, F0 magnitude of the FFR gets larger as you increase pitch salience, and it is similarly modulated by language experience (C>E). A direct comparison of cortical and brainstem responses reveals different patterns of relative changes in magnitude along the pitch salience continuum (FFR, monotonic; CPR, non-monotonic). These differences in sensitivity to pitch salience at the brainstem and cortical level may implicate a transformation in pitch processing at the cortical level presumably mediated by local sensory and/or extrasensory influence overlaid on the brainstem output.

3.1. Pitch relevant neural activity in the brainstem and auditory cortex is sensitive to pitch salience

Our results show that with increasing pitch salience of IRN stimuli, there is greater FFR magnitude at F0 and greater CPR magnitude with shorter latency for components Na–Pb and Pb–Nb regardless of language experience. The increase in pitch relevant neural activity in the FFR suggests that an increase in the neural periodicity strength results from increasing temporal regularity of the stimulus (Krishnan et al., 2010a, 2010b). This interpretation is consistent with perceptual (Patterson et al., 1996; Yost, 1996a) and physiologic (Sayles and Winter, 2007; Shofner, 1999) data indicating that the pitch of static and dynamic IRN stimuli is based on an autocorrelation-like temporal processing. Cariani and Delgutte (1996a) indeed observed a strong correspondence between the neural pitch strength of complex sounds and their pitch salience for auditory nerve responses.

At the cortical level, our CPR results are consistent with extant data showing that the latency and amplitude of the cortical POR varies systematically with the pitch salience of an IRN stimulus (Krumbholz et al., 2003; Seither-Preisler et al., 2006; Soeta et al., 2005). The Na component of the CPR—equivalent to

the POR—reflects neural activity synchronized to pitch onset and represents the integration of pitch information across frequency channels and/or the calculation of the initial pitch value and pitch strength in Heschl's gyrus (Gutschalk et al., 2004). The fact that CPR components Pb and Nb also changed with pitch salience suggest that neural activity relevant to pitch salience is also preserved at several levels of processing along the hierarchy. Source analyses of MEG derived POR (Gutschalk et al., 2002, 2004), EEG derived pitch onset responses (Bidelman and Grall, 2014), and human depth-electrode recordings (Schonwiesner and Zatorre, 2008) all indicate that Na is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing (Johnsrude et al., 2000; Zatorre, 1988). Moreover, both functional brain imaging (Griffiths et al., 1998, 2001) and intracranial recordings (Schonwiesner and Zatorre, 2008) reveal, respectively, an increase in neural activity of the primary auditory cortex and an increase in discharge rates of auditory cortical neurons as a function of iteration steps in humans.

The shortening of CPR response latency with increasing pitch salience may reflect shortening of the temporal integration window due to improved neural synchrony with increasing temporal regularity of the IRN stimuli. Whether pitch-relevant information extracted by these cortical neural generators is based on a spectral and/or temporal code is unclear. There is evidence that neurons in primary auditory cortex exhibit temporal and spectral response properties that could enable these pitch-encoding schemes (Lu et al., 2001; Steinschneider et al., 1998). However, in unanesthetized cats and primates, neurons appear to encode temporal structure using firing rate rather than using the temporal structure of the response (Dong et al., 2011; Lu et al., 2001; Wang, 2007; Wang et al., 2008; Yin et al., 2011). Thus, 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 or by a hybrid mechanism that utilizes both spectral and temporal information (Bartlett and Wang, 2007; Batra et al., 1989; Cariani and Delgutte, 1996a, 1996b; Cedolin and Delgutte, 2005; Langner, 1992), the most commonly observed code for periodicity and pitch within cortical neurons is a modulation of spike rates as a function of F0 (Bendor et al., 2012; Wang et al., 2008). 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 et al., 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. One possibility is that the temporal code is transformed in to a response synchrony code where temporally coherent activity from the subcortical stages will produce greater spike rate, yielding larger response amplitude at the cortical level. Gao and Wehr (2015) have suggested that synaptic inputs to rate-coding neurons arise in part from temporal-coding neurons, but are transformed by excitatory-inhibitory interactions. Our findings suggest that the fundamental neural mechanisms of pitch at the brainstem level and at early stages in the auditory cortex are the same for Chinese and English listeners alike. Chinese listeners, however, are more sensitive to perceptually-relevant pitch

attributes by virtue of their long-term experience with a tonal language (Gandour and Krishnan, 2014, 2016).

3.2. Is there a transformation of neural representation of pitch salience from the brainstem to auditory cortex?

A direct comparison of changes in magnitude of pitch-relevant neural activity between successive iteration steps along the pitch salience continuum revealed different patterns of relative changes in magnitude for brainstem and cortical responses (FFR, monotonic; CPR, non-monotonic). This raises the question whether these differences simply reflect normal variations in amplitude growth functions at the two levels or alternatively, fundamental differences in the sensitivity to changes in pitch salience at the two levels. The latter interpretation would implicate a transformation in pitch processing at the cortical level presumably mediated by local sensory and/or extrasensory influence overlaid on the brainstem output.

Based on the strong correlation in neural activity relevant to pitch salience between the brainstem and the auditory cortex recorded concurrently (Krishnan et al., 2012a), we concluded that brainstem and cortical representations of pitch are shaped in a coordinated manner through cortico-fugal modulation of subcortical afferent circuitry. The reverse hierarchy theory provides a representational hierarchical framework to describe the interaction between sensory input and top-down processes in primary sensory areas (Ahissar and Hochstein, 2004; Nahum et al., 2008). Its basic claim is 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 fine grained high resolution information to optimize perception. This claim is supported by neuroanatomical evidence for ascending and descending pathways (Saldana et al., 1996) and physiologic evidence for cortical modulation of brainstem representations (Suga and Ma, 2003). In the case of humans, the reverse hierarchy theory has been invoked to explain top-down enhancement of brainstem pitch representations that result from short-term auditory training (Song et al., 2008); long-term linguistic experience (Krishnan et al., 2012b); and musical training (Musacchia et al., 2008). Consistent with this theory, brainstem representation of spectrotemporal features of pitch would be more fine-grained compared to early coarse-grained, sensory representations in the auditory cortex. Indeed, the latter have been shown to be more labile and spatiotemporally broader (Chechik et al., 2006; Warren and Griffiths, 2003; Winer et al., 2005; Zatorre and Belin, 2001). In this study, we infer that differences between these two levels of brain structure in terms of the pattern of relative changes in magnitude along the pitch salience continuum implicate a transformation in the nature of pitch processing.

3.3. Language experience enhances brainstem and cortical representation of information relevant to pitch salience

Our findings show a language-dependent response enhancement for Chinese listeners in both the brainstem FFR component at F0 and the cortical CPR component Na–Pb across stimuli varying in pitch salience. The language-dependent

enhancement of the brainstem FFR component is in agreement with an earlier report on IRN stimuli varying in pitch salience (Krishnan et al., 2010c). Given that the pattern of amplitude change with pitch salience is similar for both groups, we infer that extrasensory processes are overlaid on sensory processes to modulate long-term, experience-driven, adaptive pitch mechanisms in the brainstem and at early sensory levels of pitch processing in the auditory cortex. This is accomplished by sharpening response properties of neural elements to enable optimal representation of behaviorally relevant dimensions of pitch over a broader dynamic range of pitch salience. That is, larger amplitude and shorter latency for the Chinese cortical responses may reflect the robustness of the underlying pitch-relevant neural activity and shorter integration times within temporal windows being utilized to process the various temporal attributes of pitch. Because enhanced sensitivity to pitch salience is already present in neural activity at the level of the brainstem (Krishnan et al., 2010a), we propose that cortical pitch mechanisms may be reflecting, at least in part, this enhanced pitch input from the brainstem (Krishnan et al., 2012a). This interpretation is consistent with previous studies of short-term auditory training (Russo et al., 2005; Song et al., 2008), long-term language experience (Bidelman et al., 2011, 2013, 2014); and musical training (Bidelman and Alain, 2015; Musacchia et al., 2007; Wong et al., 2007).

It is unclear why the experience-dependent enhancement was observed only for Na–Pb and not for Pb–Nb. In previous reports, we manipulated either the shape of pitch contours (Krishnan et al., 2015a, 2015b) or pitch acceleration (Krishnan et al., 2015c), and consistently observed experience-dependent enhancement of response magnitude for both Na–Pb and Pb–Nb. One plausible explanation is that neural activity in the Na–Pb time window optimally represents neural processing relevant to pitch salience, and therefore would be subject to experience-dependent modulation. The time window for Pb–Nb, on the other hand, may be indexing other dynamic attributes of pitch: e.g., shape of pitch contour (Krishnan et al., 2015a, 2015b) and pitch acceleration (Krishnan et al., 2015c). 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, e.g., in this study, pitch salience. More generally, 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.

As indexed by Na–Pb amplitude over the temporal electrode sites (T7/T8), a rightward asymmetry is limited to the Chinese group for all stimuli along the pitch salience continuum. It is only over the right temporal site that the amplitude of these components are larger for the Chinese compared to English. Our findings converge with an extant literature that supports the role of the right hemisphere in processing linguistic as well as nonlinguistic pitch (Friederici and Alter, 2004; Friederici, 2011; Meyer, 2008; Zatorre and Gandour, 2008). The superiority of the Chinese group demonstrates that extrasensory components may mask purely sensory effects in their influence within a given temporal integration window. Our failure to observe language dependent asymmetry at the F3/F4 electrode sites—consistent with our previous findings (Krishnan et al., 2015a, 2015b, 2015c)—suggests that the temporal electrodes proximal to and located over the auditory cortices, the putative regions for pitch processing, are better situated to capture the experience-dependent, preferential recruitment of the right auditory cortex for pitch processing.

Thus, CPR components may capture both experience-dependent extrasensory influences as well experience-independent sensory effects (Krishnan et al., 2015b). It is true that speakers of all languages have some experience in pitch processing. In English, for example, pitch variations may signal multisyllabic word- or phrase-level stress or sentence-level intonation contrasts. Mandarin, on the other hand, exploits variations in pitch at the level of a mono-syllable to signal changes in word meaning. 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 et al., 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.

3.4. Predictive coding may underlie experience-dependent processing of pitch

Neural processes mediating language experience-dependent shaping of subcortical and cortical stages of pitch processing likely involve a coordinated interplay between bottom-up, top-down, and local neural pathways that engage both sensory and extrasensory components. This interplay is essential to extract optimal early representations of stimulus dimensions that transform, later functionally more salient cortical representations that drive processes mediating linguistic performance. Motivated by the hierarchical processing framework for cortical pitch processing (Kumar and Schonwiesner, 2012), we present here an expanded integrated predictive coding framework that adds the brainstem component to the originally proposed cortical pitch processing hierarchy in an effort to account for the experience-dependent effects observed in this study. At the cortical level, a hierarchical processing framework for coordinated interaction between both the primary auditory cortex in the medial Heschl's gyrus (HG) as well as in the adjacent more lateral non-primary areas in the HG is provided by applying a predictive coding model of perception to depth-electrode recordings of pitch-relevant neural activity along HG (Kumar et al., 2011; Kumar and Schonwiesner, 2012; cf. Rao and Ballard, 1999). Essentially, higher-level areas in the

hierarchy contributing to pitch (lateral HG) use stored information of pitch to make an initial pitch prediction. Relevant here is Hasson et al. (2015) process memory framework described in Section 3.3. Specifically, process memory is engaged in the prediction process wherein cortical circuits utilize active traces of past information that are used by a neural circuit to process incoming information in the present moment to generate a prediction. Given the rapid neural dynamics required, the receptive windows would have to be necessarily short for processing the dynamic pitch contours used in this study. This prediction is passed to the lower areas in the processing hierarchy at the cortical level (medial and middle HG) via top down connection(s), and to subcortical level (inferior colliculus (IC), the presumed source of FFR) via the corticofugal pathways. The lower areas at the cortical level, and at IC then compute a prediction error. The strength of the top-down and bottom-up connections to each level is continually adjusted in a recursive manner in order to minimize predictive error at each level in the hierarchy to optimize representation at the higher level. Consistent with the predictions of the model, Kumar et al. (2011) showed that strength of connectivity at the cortical level varies with pitch salience such that the strength of the top down connection from lateral HG to medial and middle HG increased with pitch salience, whereas the strength of the bottom up connection from middle HG to lateral HG decreased. The lateral HG has more pitch-specific mechanisms, and therefore plays a relatively greater role in pitch perception. It is likely that similar changes in the strength of the corticofugal inputs to the IC with changes in pitch salience may account for the experience-dependent enhancement of the subcortical responses. It has been proposed that the corticofugal inputs to the IC provide continuous online modulation of processing of pitch-relevant information based on a predictive algorithm wherein robust representation of behaviorally relevant features suggest smaller predictive error (Chandrasekaran et al., 2014).

Applied to our data, this framework suggests that CPR (cortical) and FFR (brainstem) 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). The hierarchical flow of processing and its connectivity strengths along the HG, and to the IC are essentially the same regardless of one's language background. However, the initial pitch prediction at cortical and subcortical levels is more precise for Chinese because of their access to stored information about native pitch contour (T2) with a smaller error term. Consequently, the top-down connections at both levels (lateral HG to medial and middle HG at the cortical level; and corticofugal to the IC) are 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 at both cortical and subcortical levels to optimally extract behaviorally relevant features of sounds.

Pitch processing in the auditory cortex is also influenced by inputs from subcortical structures that are also subject to experience-dependent plasticity. As in the cortical level

hierarchy, it is likely that corticofugal projections from the auditory cortex to the inferior colliculus—the presumed site of FFR generation—provide feedback to adjust the effective integration time scales at each stage of hierarchical processing to optimally control the temporal dynamics of pitch processing (Balaguer-Ballester et al., 2009). Language-dependent enhancement in the neural activity relevant to pitch salience at the brainstem level (FFR) and at the cortical level (CPR) in the 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. Conclusions

Parametric variation of pitch salience enables us to disentangle pitch-relevant neural activity that reflects primarily language universal (acoustic) sensory processes and overlaid language-dependent (linguistic) neural activity. Enhanced sensitivity to this pitch dimension at both the brainstem and early cortical sensory level of processing, as well as the strong rightward asymmetry of the Na–Pb component in the Chinese group, is consistent with the notion that long-term experience shapes adaptive, distributed hierarchical pitch processing in the auditory cortex, and reflects an interaction with higher-order, extrasensory processes beyond the sensory memory trace. The restriction to Na–Pb suggests that the relative weighting of CPR components vary depending on the sensitivity of neural activity within a particular temporal window to a specific attribute of pitch. Differences in sensitivity to pitch salience at the brainstem and cortical level may implicate a transformation in the nature of pitch processing at the cortical level.

5. Experimental procedures

5.1. Participants

EEG data were recorded from a total of seventeen native speakers of Mandarin Chinese (7 male, 10 female) and English (8 male, 9 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 closely matched in age (Chinese: 23.9 ± 3.7 years; English: 23.6 ± 2.9), years of formal education (Chinese: 16.88 ± 2.15 years; English: 15.9 ± 1.3), and were strongly right handed (Chinese: $90.7 \pm 14.2\%$; English: 96.0 ± 8.2) 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 (10.7 ± 1.4 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, 4.8 and 5.3, respectively (Li et al., 2006). Their daily usage of Mandarin and English, in order, was

reported to be 65% and 35%. As determined by a music history questionnaire (Wong and Perrachione, 2007), all Chinese and English participants had less than two years of musical training (Chinese, 0.59±0.88 years; English, 1.12±1.73) 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.

5.2.   Stimuli

Four IRN stimuli, each with a pitch contour exemplary of Mandarin T2 (cf. Krishnan et al., 2010c; Fig. 3, A1), were designed to represent a pitch salience continuum ranging from weak pitch (I4) to strong pitch (I32) and two intermediate levels of pitch salience—one toward weak (I8), and another toward strong (I16). Iterated rippled noise was used to create these stimuli by applying polynomial equations that generate dynamic, curvilinear pitch patterns (Swaminathan et al., 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 (80–4000 Hz) and adding it back on itself in a recursive manner (Yost, 1996b). The pitch of IRN corresponds to the reciprocal of the delay (1/d); its salience grows with the number of iterations (Krishnan et al., 2010a; Patterson et al., 1996; Yost, 1996b) with little or no change in salience beyond an iteration step of 32 (Yost, 1996a), the upper limit used here. Thus, IRN is a complex pitch-evoking stimulus, which allows one to systematically manipulate the temporal regularity and hence pitch salience of a stimulus. These four IRN stimuli were used in an experimental paradigm which consisted of three segments: a 250 ms pitch segment (I4, I8, I16, I32) preceded by a 750 ms noise segment and followed by a 235 ms noise segment (Fig. 9, A). The noise segments were crossfaded with the pitch segment using 7.5 ms $\cos^2$ ramps.
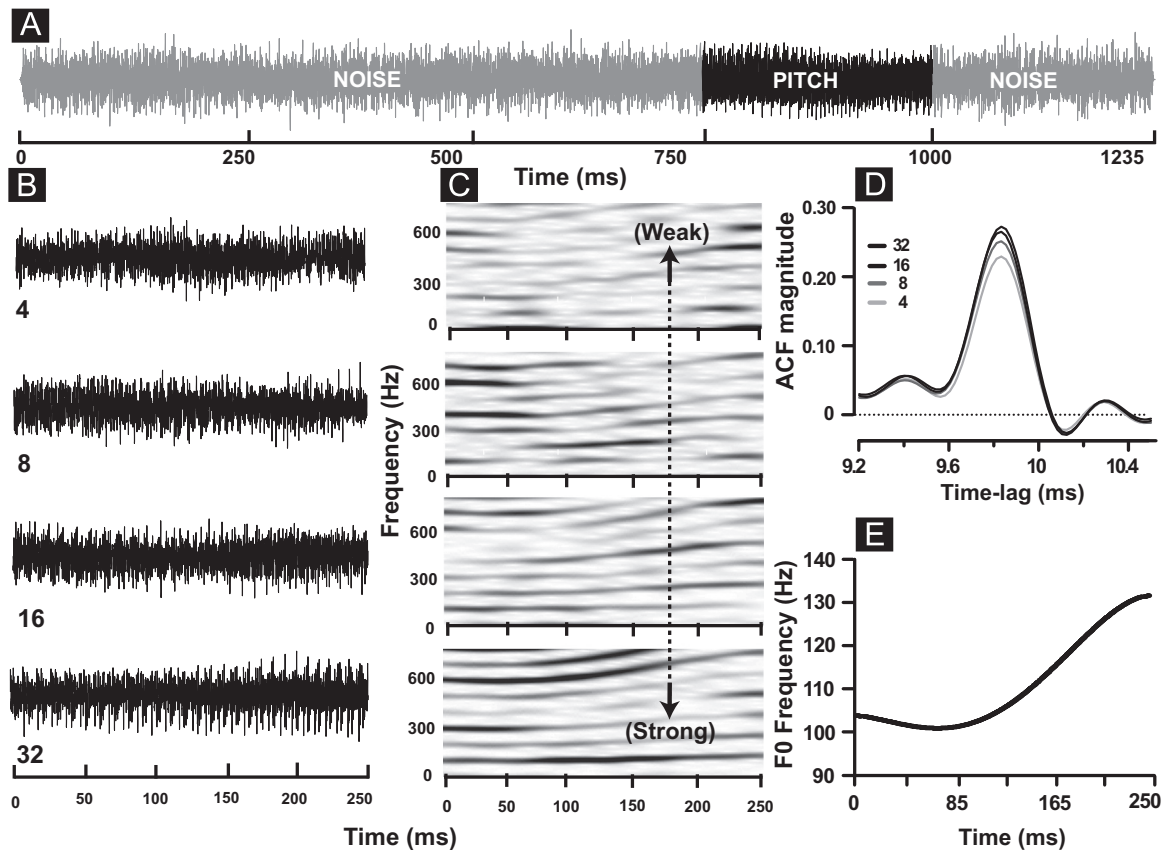

Fig. 9 – Stimulus paradigm used to acquire cortical and brainstem responses to Mandarin Tone 2 varying in degrees of pitch salience (A). The waveform shows robust periodicity in the pitch segment at a strong iteration step (n=32). Waveforms (B), spectrograms (C), and stimulus ACFs (D) for weak (n=4), medium weak (n=8), medium strong (n=16), and strong (n=32), where n=number of iteration steps. Waveforms and spectrograms, respectively, reveal more robust periodicity and spectral information with increasing n. Normalized magnitude of spectrograms are represented using the gray scale gradient. Darker shades indicate larger spectral magnitudes. ACF magnitude around the fundamental pitch period of the stimuli (F0=101.5 Hz; period=9.85 ms) increases steadily as the number of iterations (n) is increased, indicating that acoustic periodicity improves with increasing n. The pitch salience of the voice fundamental frequency contour of Tone 2 (E) is varied by using a different number of iterations (n) in the IRN generating circuit. ACF, autocorrelation function; F0, fundamental frequency; IRN, iterated rippled noise.

Temporal and spectral characteristics of the pitch stimuli are displayed in Fig. 9 (B–E).

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 0.56/s. Stimulus presentation order was randomized both within and across participants. The overall root-mean-square level of each segment was equated such that there was no discernible difference in intensity among the three segments. All stimuli were generated and played out using an auditory evoked potential system (SmartEP, Intelligent Hearing Systems; Miami, FL, USA).

5.3. Cortical and brainstem 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; 0.3–2500 Hz analog band-pass) through the ASA-Lab EEG 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. 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 characterize the 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 kΩ. For each stimulus, EEGs were acquired in two blocks of 1000 sweeps each. The experimental protocol took about 2 h to complete.

5.3.1. Extraction of CPR and FFR

CPR and FFR responses were extracted off-line from the EEG files using different re-referenced electrode montages. We first analyzed the CPR data on fourteen participants in each group (Chinese, 7 male, 7 female; English, 6 male, 8 female). Subsequent FFR analysis on the same EEG data revealed noisy FFR responses from three participants in each group that were not amenable to analysis. They were replaced by FFR data from three new participants for each group. This new data set ($n=14$) was used to characterize the FFR responses only. A reduced data set ($n=11$) was analyzed to ensure that CPR/FFR comparisons were made using data from the same subjects from each language group.

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 $\pm50\,\mu V$ 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 cross-language 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 (Krishnan et al., 2014a, 2015a, 2015b). 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 fronto-central sites (e.g., Bidelman and Grall, 2014; Krishnan et al., 2015b; Krumbholz et al., 2003). It also allows us to compare our CPR data with Fz-derived POR data (Bidelman and Grall, 2014; Gutschalk et al., 2002, 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.

To extract the FFR, EEG files were digitally band-pass filtered (75–1500 Hz, Butterworth zero phase filters with 24 dB/Octave rejection rate). Sweeps containing electrical activity exceeding $\pm40\,\mu V$ were rejected automatically. Subsequently time domain averaging was performed on three different re-referenced electrode montages (FPz-linked mastoids; Fz-linked mastoids; and Cz-linked mastoids) over an analysis window of 270 ms (from 743 to 1013 ms, where 743 represents the onset of the pitch segment). Each FFR waveform represents the grand average of the FFRs derived from the three electrode montages to a total of 2000 sweeps presented in 2 blocks of 1000 sweeps each. These three channels were chosen because of the prominence of the FFRs in fronto-central locations—typical configurations used to record FFRs. The rationale for averaging across channels was to improve detectability of the FFR, particularly for stimuli with weaker pitch salience by enhancing the robustness of FFR signals, and some reduction of noise resulting from averaging.

5.3.2. CPR latency and 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:125–150 ms, Pb: 200–220,

Nb:270–285 ms) 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 (Krishnan et al., 2014a, 2014b). We evaluated only the latency and magnitude of the CPR components. Peak latencies of Na, Pb, and Nb (time interval between pitch-eliciting stimulus onset and response peak of interest) and peak-to-peak amplitude of Na–Pb and Pb–Nb were measured manually to characterize the effects of changes in pitch salience throughout the temporal course of the pitch contours. Manual peak picking was preferred due to decreased reliability of automatic peak picking associated with small and degraded response waveform morphology. To improve accuracy and consistency of manual peak picking for latency and amplitude measurements, individual averaged responses were overlaid on the grand averaged response to facilitate response identification. The time point around the peak with maximum voltage was taken as the measure of absolute latency and peak amplitude for a given component. This process was repeated independently by two members in the laboratory, exhibiting high interjudge reliability (90%). For each condition, 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 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 frontal and temporal electrodes.

5.3.3. FFR neural pitch strength

Neural pitch strength was quantified by measuring the magnitude of the F0 component from each response waveform. The spectrum of each response segment was computed by taking the Fast Fourier Transform (FFT) of a time-windowed version of its temporal waveform (Gaussian window, 1 Hz resolution). For each subject, the magnitude of F0 was measured as the peak in the FFT, relative to the noise floor. All FFR data analyses were performed using custom routines coded in Matlab 11 (The MathWorks, Inc., Natick, MA, USA).

5.3.4. Comparison of CPR and FFR

Comparison of the CPR and FFR data was limited to the original 11 (Chinese: 6 male, 5 female; English: 5 male, 6 female) out of 14 participants for whom we were able to analyze both CPR and FFR data. CPR components are several orders of magnitude larger than the FFR response. The magnitudes of both cortical and brainstem responses were therefore expressed as a normalized ratio between successive iteration steps ((I8-I4)/I32; (I16-I8)/I32; (I32–I16)/I32), referred to as I8–I4, I16–I8, and I32–I16, respectively. This procedure allowed for a meaningful comparison of the change in pitch-related neural activity at the brainstem and cortical levels with changes in pitch salience, independent of their differences in absolute magnitude.

5.4. Statistical analysis

Separate, two-way (group × stimulus), mixed model ANOVAs (SAS®; SAS Institute, Inc., Cary, NC, USA) were performed on each component of peak latency (Na, Pb, and Nb) and peak-to-peak amplitude (Na–Pb, Pb–Nb) derived from the Fz electrode site; a three-way (group × stimulus × 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 (I4, I8, I16, I32) and electrode site (T7 [left]/T8 [right]) were treated as within-subject factors. A two-way (group × stimulus) ANOVA was performed on neural pitch strength derived from FFRs to assess how subcortical encoding of pitch-related information varies as a function of pitch salience. To compare the magnitude of cortical (Na–Pb, Pb–Nb) and brainstem (FFR) response components, separate three-way ANOVAs (component × group × stimulus) were performed on ratios derived from paired iteration steps. Stimulus (I4–I8, I8–I16, I16–I32) and component (Na–Pb, FFR; Pb–Nb, FFR) 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 (η_p^2) values, where appropriate, were reported to indicate effect sizes.

Appendix A.   Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.brainres.2016.02.021.

REFERENCES

Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. 8, 457–464.

Balaguer-Ballester, E., Clark, N.R., Coath, M., Krumbholz, K., Denham, S.L., 2009. Understanding pitch perception as a hierarchical process with top-down modulation. PLoS Comput. Biol. 5, e1000301.

Bartlett, E.L., Wang, X., 2007. Neural representations of temporally modulated signals in the auditory thalamus of awake primates. J. Neurophysiol. 97, 1005–1017.

Batra, R., Kuwada, S., Stanford, T.R., 1989. Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. J. Neurophysiol. 61, 257–268.

Bendor, D., Osmanski, M.S., Wang, X., 2012. Dual-pitch processing mechanisms in primate auditory cortex. J. Neurosci. 32, 16149–16161.

Bidelman, G.M., Gandour, J.T., Krishnan, A., 2011. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J. Cogn. Neurosci. 23, 425–434.

Bidelman, G.M., Krishnan, A., 2011. Brainstem correlates of behavioral and compositional preferences of musical harmony. Neuroreport 22, 212–216.

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

Bidelman, G.M., Grall, J., 2014. Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. Neuroimage 101, 204–214.

Bidelman, G.M., Weiss, M.W., Moreno, S., Alain, C., 2014. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. Eur. J. Neurosci. 40, 2662–2673.

Bidelman, G.M., Alain, C., 2015. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. J. Neurosci. 35, 1240–1249.

Cariani, P.A., Delgutte, B., 1996a. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J. Neurophysiol. 76, 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. J. Neurophysiol. 76, 1717–1734.

Cedolin, L., Delgutte, B., 2005. Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. J. Neurophysiol. 94, 347–362.

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

Chandrasekaran, B., Skoe, E., Kraus, N., 2014. An integrative model of subcortical auditory plasticity. Brain Topogr. 27, 539–552.

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, 359–368.

Cowan, N., 1988. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychol. Bull. 104, 163–191.

Dong, C., Qin, L., Liu, Y., Zhang, X., Sato, Y., 2011. Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. PLoS One 6, e25895.

Friederici, A.D., Alter, K., 2004. Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang. 89, 267–276.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91, 1357–1392.

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

Gandour, J.T., Krishnan, A., 2016. Processing tone languages. In: Hickok, G., Small, S.L. (Eds.), Neurobiology of Language. Academic Press, New York, pp. 1095–1107.

Gao, X., Wehr, M., 2015. A coding transformation for temporally structured sounds within auditory cortical neurons. Neuron 86, 292–303.

Griffiths, T.D., Buchel, C., Frackowiak, R.S., Patterson, R.D., 1998. Analysis of temporal structure in sound by the human brain. Nat. Neurosci. 1, 422–427.

Griffiths, T.D., Uppenkamp, S., Johnsrude, I., Josephs, O., Patterson, R.D., 2001. Encoding of the temporal regularity of sound in the human brainstem. Nat. Neurosci. 4, 633–637.

Griffiths, T.D., Kumar, S., Sedley, W., Nourski, K.V., Kawasaki, H., Oya, H., Patterson, R.D., Brugge, J.F., Howard, M.A., 2010. Direct recordings of pitch responses from human auditory cortex. Curr. Biol. 20, 1128–1132.

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, 207–216.

Gutschalk, A., Patterson, R.D., Scherg, M., Uppenkamp, S., Rupp, A., 2004. Temporal dynamics of pitch in human auditory cortex. Neuroimage 22, 755–766.

Hasson, U., Chen, J., Honey, C.J., 2015. Hierarchical process memory: memory as an integral component of information processing. Trends Cogn. Sci. 19, 304–313.

Herdman, A.T., Lins, O., Van Roon, P., Stapells, D.R., Scherg, M., Picton, T.W., 2002. Intracerebral sources of human auditory steady-state responses. Brain Topogr. 15, 69–86.

Johnsrude, I.S., Penhune, V.B., Zatorre, R.J., 2000. Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123, 155–163.

Krishnan, A., 2007. Human frequency following response. In: Burkard, R.F., Don, M., Eggermont, J.J. (Eds.), Auditory Evoked Potentials: Basic Principles and Clinical Application. Lippincott Williams & Wilkins, Baltimore, pp. 313–335.

Krishnan, A., Plack, C.J., 2009. Auditory brainstem correlates of basilar membrane nonlinearity in humans. Audiol. Neurootol. 14, 88–97.

Krishnan, A., Bidelman, G.M., Gandour, J.T., 2010a. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. Hear. Res. 268, 60–66.

Krishnan, A., Gandour, J.T., Bidelman, G.M., 2010b. Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. Brain Res. 1313, 124–133.

Krishnan, A., Gandour, J.T., Smalt, C.J., Bidelman, G.M., 2010c. Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. Brain Lang. 114, 193–198.

Krishnan, A., Bidelman, G.M., Smalt, C.J., Ananthakrishnan, S., Gandour, J.T., 2012a. Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. Neuropsychologia 50, 2849–2859.

Krishnan, A., Gandour, J.T., Bidelman, G.M., 2012b. Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. Neuroreport 23, 498–502.

Krishnan, A., Gandour, J.T., 2014. Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: electrophysiological evidence. Acoust. Aust. 42, 166–178.

Krishnan, A., Gandour, J.T., Ananthakrishnan, S., Vijayaraghavan, V., 2014a. Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. Neuropsychologia 59, 1–12.

Krishnan, A., Gandour, J.T., Suresh, C.H., 2014b. Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. Brain Lang. 138, 51–60.

Krishnan, A., Gandour, J.T., Ananthakrishnan, S., Vijayaraghavan, V., 2015a. Language experience enhances early cortical pitch-dependent responses. J. Neurolinguistics 33, 128–148.

Krishnan, A., Gandour, J.T., Suresh, C.H., 2015b. Pitch processing of dynamic lexical tones in the auditory cortex is influenced by sensory and extrasensory processes. Eur. J. Neurosci. 41, 1496–1504.

Krishnan, A., Gandour, J.T., Suresh, C.H., 2015c. Experience-dependent enhancement of pitch-specific responses in the auditory cortex is limited to acceleration rates in normal voice range. Neuroscience 303, 433–445.

Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lutkenhoner, B., 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb. Cortex 13, 765–772.

Kumar, S., Sedley, W., Nourski, K.V., Kawasaki, H., Oya, H., Patterson, R.D., Howard 3rd, M.A., Friston, K.J., Griffiths, T.D., 2011. Predictive coding and pitch processing in the auditory cortex. J. Cogn. Neurosci. 23, 3084–3094.

Kumar, S., Schonwiesner, M., 2012. Mapping human pitch representation in a distributed system using depth-electrode recordings and modeling. J. Neurosci. 32, 13348–13351.

Langner, G., 1992. Periodicity coding in the auditory system. Hear. Res. 60, 115–142.

Li, P., Sepanski, S., Zhao, X., 2006. Language history questionnaire: a web-based interface for bilingual research. Behav. Res. Methods 38, 202–210.

Lu, T., Liang, L., Wang, X., 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat. Neurosci. 4, 1131–1138.

Meyer, M., 2008. Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. Z. Neuropsychol. 19, 101–115.

Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc. Natl. Acad. Sci. USA 104, 15894–15898.

Musacchia, G., Strait, D., Kraus, N., 2008. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear. Res. 241, 34–42.

Nahum, M., Nelken, I., Ahissar, M., 2008. Low-level information and high-level perception: the case of speech in noise. PLoS Biol. 6, e126.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. Ear Hear. 30, 653–661.

Patel, A.D., 2008. Music, Language, and the Brain. Oxford University Press, NY.

Patterson, R.D., Handel, S., Yost, W.A., Datta, A.J., 1996. The relative strength of the tone and noise components in iterated ripple noise. J. Acoust. Soc. Am. 100, 3286–3294.

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. J. Neurosci. 24, 6810–6815.

Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.

Russo, N.M., Nicol, T.G., Zecker, S.G., Hayes, E.A., Kraus, N., 2005. Auditory training improves neural timing in the human brainstem. Behav. Brain Res. 156, 95–103.

Saldana, E., Feliciano, M., Mugnaini, E., 1996. Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. J. Comp. Neurol. 371, 15–40.

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 Res. 1171, 52–66.

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. Exp. Brain Res. 187, 97–105.

Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., Lutkenhoner, B., 2006. Evidence of pitch processing in the N 100 m component of the auditory evoked field. Hear. Res. 213, 88–98.

Shofner, W.P., 1999. Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural auto-correlograms. J. Neurophysiol. 81, 2662–2674.

Shofner, W.P., 2002. Perception of the periodicity strength of complex sounds by the chinchilla. Hear. Res. 173, 69–81.

Soeta, Y., Nakagawa, S., Tonoike, M., 2005. Auditory evoked magnetic fields in relation to iterated rippled noise. Hear. Res. 205, 256–261.

Song, J.H., Skoe, E., Wong, P.C.M., Kraus, N., 2008. Plasticity in the adult human auditory brainstem following short-term linguistic training. J. Cogn. Neurosci. 20, 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. J. Acoust. Soc. Am. 104, 2935–2955.

Steinschneider, M., Volkov, I.O., Noh, M.D., Garell, P.C., Howard 3rd, M.A., 1999. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. J. Neurophysiol. 82, 2346–2357.

Suga, N., Ma, X., 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. Nat. Rev. Neurosci. 4, 783–794.

Suga, N., Ma, X., Gao, E., Sakai, M., Chowdhury, S.A., 2003. Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. Speech Commun. 41, 189–200.

Swaminathan, J., Krishnan, A., Gandour, J.T., 2008. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. Neuroreport 19, 1163–1167.

Walker, K.M., Bizley, J.K., King, A.J., Schnupp, J.W., 2011. Cortical encoding of pitch: recent results and open questions. Hear. Res. 271, 74–87.

Wang, X., 2007. Neural coding strategies in auditory cortex. Hear. Res. 229, 81–93.

Wang, X., Lu, T., Bendor, D., Bartlett, E., 2008. Neural coding of temporal information in auditory thalamus and cortex. Neuroscience 157, 484–494.

Warren, J.D., Griffiths, T.D., 2003. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J. Neurosci. 23, 5799–5804.

Winer, J.A., Miller, L.M., Lee, C.C., Schreiner, C.E., 2005. Auditory thalamocortical transformation: structure and function. Trends Neurosci. 28, 255–263.

Wong, P.C., Perrachione, T.K., 2007. Learning pitch patterns in lexical identification by native English-speaking adults. Appl. Psycholinguist. 28, 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. Nat. Neurosci. 10, 420–422.

Xu, Y., Gandour, J.T., Francis, A.L., 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. J. Acoust. Soc. Am. 120, 1063–1074.

Yin, P., Johnson, J.S., O'Connor, K.N., Sutter, M.L., 2011. Coding of amplitude modulation in primary auditory cortex. J. Neurophysiol. 105, 582–600.

Yost, W.A., 1996a. Pitch strength of iterated rippled noise. J. Acoust. Soc. Am. 100, 3329–3335.

Yost, W.A., 1996b. Pitch of iterated rippled noise. J. Acoust. Soc. Am. 100, 511–518.

Zatorre, R.J., 1988. Pitch perception of complex tones and human temporal-lobe function. J. Acoust. Soc. Am. 84, 566–572.

Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11, 946–953.

Zatorre, R.J., Gandour, J.T., 2008. Neural specializations for speech and pitch: moving beyond the dichotomies. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 1087–1104.

Zatorre, R.J., Baum, S.R., 2012. Musical melody and speech intonation: singing a different tune. PLoS Biol. 10, e1001372.