



Published in final edited form as:

Neuroscience. 2017 March 27; 346: 52–63. doi:10.1016/j.neuroscience.2017.01.013.

Changes in pitch height elicit both language universal and language dependent changes in neural representation of pitch in the brainstem and auditory cortex

Ananthanarayan Krishnan^a, Chandan H. Suresh^a, and Jackson T. Gandour^a

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

Abstract

Language experience shapes encoding of pitch-relevant information at both brainstem and cortical levels of processing. Pitch height is a salient dimension that orders pitch from low to high. Herein we investigate the effects of language experience (Chinese, English) in the brainstem and cortex on i) neural responses to variations in pitch height, ii) presence of asymmetry in cortical pitch representation, and iii) patterns of relative changes in magnitude of pitch height between these two levels of brain structure. Stimuli were three nonspeech homologs of Mandarin Tone 2 varying in pitch height only. The frequency-following response (FFR) and the cortical pitch-specific response (CPR) were recorded concurrently. At the Fz-linked T7/T8 site, peak latency of Na, Pb, and Nb decreased with increasing pitch height for both groups. Peak-to-peak amplitude of Na–Pb and Pb–Nb increased with increasing pitch height across groups. A language-dependent effect was restricted to Na–Pb; the Chinese had larger amplitude than the English group. At temporal sites (T7/T8), the Chinese group had larger amplitude, as compared to English, across stimuli, but also limited to the Na–Pb component and right temporal site. In the brainstem, F0 magnitude decreased with increasing pitch height; Chinese had larger magnitude across stimuli. A comparison of CPR and FFR responses revealed distinct patterns of relative changes in magnitude common to both groups. CPR amplitude increased and FFR amplitude decreased with increasing pitch height. Experience-dependent effects on CPR components vary as a function of neural sensitivity to pitch height within a particular temporal window (Na–Pb). Differences between the auditory brainstem and cortex imply distinct neural mechanisms for pitch extraction at both levels of brain structure.

Keywords

auditory; pitch encoding; pitch height; iterated rippled noise; cortical pitch response; fundamental frequency response

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

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

INTRODUCTION

Pitch is a multidimensional perceptual attribute that plays an important role in speech, language and hearing. One salient dimension of pitch is pitch height, which orders pitch from low to high. Perceptually, pitch height is a continuous vertical dimension associated with the voice fundamental frequency (F0) of sound that varies directly with frequency, and provides a basis for segregation of sound sources (Roffler and Butler, 1968; Shepard, 1982; Krumhansl, 1990; Melara and Marks, 1990). Magnetoencephalography (MEG) studies show that shorter latencies of different auditory evoked fields (m100; pitch onset response, POR) are associated with increasing pitch of pure tones (Roberts and Poeppel, 1996; Seither-Preisler, Krumbholz and Lutkenhoner, 2003); synthetic speech (Poeppel and Roberts, 1996); bandpass-filtered harmonic tones (Ragot and Lepaul-Ercole, 1996); and iterated rippled noise stimuli (IRN) (Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner, 2003; Ritter, Dosch, Specht, Schneider and Rupp, 2007). These findings converge on the notion that integration times required for temporal pitch extraction change as a function of pitch height. With the exception of the POR response, other auditory evoked fields (e.g., m100, MMN) do not reflect neural activity specific to pitch. As revealed by functional neuroimaging, areas specifically activated by pitch height changes lie posterior to primary auditory cortex within the planum temporale (Griffiths and Warren, 2002; Warren, Uppenkamp, Patterson and Griffiths, 2003).

Pitch contour refers to the shape and direction of the trajectory shown by any perceptible change in pitch value through the duration of the syllable; pitch height refers to the relative placement of the values within the values of the pitch span (Laver, 1994, pp. 460–461). Both pitch contour and pitch height play an important role in the perception of lexical tones (for reviews, Gandour, 1994; Gandour and Krishnan, 2014, 2016). Multidimensional scaling studies, however, show that the pitch contour dimension is a better discriminator than pitch height in distinguishing tone-language native speakers from non-tone language. Whereas pitch contour is a relatively more important perceptual cue to Mandarin speakers, the opposite is true for English (Gandour, 1983; Huang and Johnson, 2011). This differential pattern of relative weighting between pitch contour and pitch height is preserved even when trained Mandarin and English speakers are asked to make pairwise difference ratings of Cantonese lexical tones (Francis, Ciocca, Ma and Fenn, 2008). Auditory evoked responses in the cerebral cortex similarly reveal a differential sensitivity to pitch contour and pitch height. Mismatch negativity (MMN) responses evoked by Mandarin tones show that Chinese listeners are more sensitive to pitch direction than height as compared to English listeners (Chandrasekaran, Gandour and Krishnan, 2007). As indexed by MMN, the time courses of these pitch dimensions also appear to be different. In Mandarin, pitch height contrasts elicit shorter peak latencies than those based on pitch contour (Wang, Wang and Chen, 2013). In Cantonese (Tsang, Jia, Huang and Chen, 2011), MMN amplitude as well as latency is shown to be modulated by the magnitude of the change in height. The location of the turning point, however, varied across pitch contours. Changes in location may have also had an effect on pitch height. Moreover, MMN is not a cortical pitch-specific response. Questions therefore remain about the latency and amplitude of pitch height.

Of particular interest to this study are cortical pitch response (CPR) data from our previous experiments. Language experience effects are not uniform across temporal attributes of pitch. To date, we have identified two temporal attributes of dynamic pitch: rate of pitch acceleration and pitch salience. At the frontocentral Fz electrode site, amplitude of both CPR components (Na–Pb, Pb–Nb) are modulated by language experience when pitch stimuli varying in acceleration rates are also accompanied by changes in other temporal pitch attributes. For example, overall trajectory (flat, linear, or curvilinear) of pitch contours (Krishnan, Gandour and Suresh, 2014; Krishnan, Gandour, Ananthkrishnan and Vijayaraghavan, 2015; Krishnan, Gandour and Suresh, 2015a); changes in location of peak acceleration (Krishnan, Gandour and Suresh, 2015b); or changes in location of turning point (Krishnan et al., submitted). However, if changes in pitch acceleration are held constant while systematically varying pitch salience across stimuli, language-dependent effects are restricted to the earlier Na–Pb time window only (Krishnan, Gandour and Suresh, 2016). This finding raises the question whether experience-dependent effects may be restricted to time windows that are optimal for processing specific temporal pitch attributes. Yet to be investigated with CPR are changes in pitch height, a temporal attribute that is associated with relative differences in the level of pitch for dynamic curvilinear pitch contours. The cortical pitch onset response (POR)—equivalent to the Na component of the CPR—has recently been shown to be sensitive to the height dimension elicited by flat (steady-state) pitch stimuli (Bidelman, 2015). POR latency occurred earlier for high (132 Hz) than low (103 Hz) flat pitch. There were no stimulus effects for amplitude. This null effect is due to the constant rate of pitch acceleration across stimuli (cf. Xu, Krishnan and Gandour, 2006; Krishnan, Gandour, Bidelman and Swaminathan, 2009; Krishnan, Gandour and Suresh, 2014). Thus, there remains a gap in our knowledge of language-dependent sensitivity to representative, time-variant pitch contours. Such stimuli must exhibit identical, changing rates of pitch acceleration, varying only in pitch height.

In this study our aim is to investigate language experience effects on representation of dynamic, time-variant pitch height information as preserved in simultaneously recorded responses in the brainstem and auditory cortex (cf. Krishnan, Bidelman, Smalt, Ananthkrishnan and Gandour, 2012, steady-state pitch). We elicited responses by systematically varying pitch height along a continuum of a time-variant IRN homolog of a Mandarin lexical tone. Concurrent recording of brainstem and cortical responses allow us to examine the interplay between subcortical and cortical levels of processing. This methodology obviates the difficulty of comparing neural responses across studies at separate levels of brain structure from different stimulation/acquisition paradigms. If we observe that patterns of relative changes in magnitude along the pitch height continuum differ between brainstem and cortical responses, we can infer that a fundamental change in the nature of processing pitch height has occurred at the cortical level. Because the stimuli are exemplary of a native Mandarin tone, we expect Chinese native speakers, as compared to English, to show enhanced sensitivity to changes in pitch height at both cortical and brainstem levels. Regarding CPR components, not all temporal attributes of pitch elicit the same language experience effects. The relative weighting of neural activity is expected to vary depending on sensitivity to a specific attribute of pitch within a particular temporal integration window. Because changes in rates of acceleration are fixed throughout stimulus duration, language-

dependent effects on pitch height (cf. pitch salience, Krishnan, Gandour and Suresh, 2016) are hypothesized to be restricted to the earlier CPR time window (Na–Pb).

EXPERIMENTAL PROCEDURES

Participants

EEG data were recorded from fourteen native speakers of Mandarin Chinese (C: 8 male, 6 female) and English (E: 5 male, 9 female), all of whom were attending Purdue University. They were matched in age (C: 25.4; E: 21.5), years of education (C: 18.5; E: 15.6), and strongly right handed (C: 95.9; E: 94.9) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971; Somers, Aukes, Ophoff, Boks, Fler, de Visser, Kahn and Sommer, 2015). All Chinese participants were born in mainland China. None had studied English before the age of nine (11.6 years). Self-ratings of English language proficiency on a Likert-type scale (1 = very poor; 7 = native-like) for speaking and listening abilities were, on average, 4.5 and 4.9, respectively (Li et al., 2006). Their daily usage of Mandarin and English, in order, were 70% and 30%. As determined by a music history questionnaire (Wong and Perrachione, 2007), all Chinese and English participants had less than two years of musical training (C: 0.56; E: 1.32) on any combination of instruments. No participant had any training within the past five years. Each participant gave written consent in accordance with the Declaration of Helsinki and with approval of the experimental protocol by the Institutional Review Board of Purdue University.

Stimuli

Fig. 1 displays voice fundamental frequency (F0) trajectories that are exemplary of the high rising Mandarin lexical tone (Xu, 1997). In this study, *changes in pitch height* does not refer to changes within stimulus, but rather overall changes in height at corresponding points in time between the three stimuli. No matter the level of pitch height, the shape and direction of the pitch contour was constant at corresponding points in time throughout the duration of the pitch span. They were designed to represent pitch height levels associated with speaker age and sex; namely, low (L), mid (M), and high (H), approximating those associated with men, women, and children, respectively (Table 1). Mean F0 of L, M, and H were 108, 198, and 288 Hz, respectively, with a fixed F0 of 90 Hz between means of both L and M and M and H. Compared to Peterson and Barney (1952), mean F0s of L and M were lower than men (132) and women (223); H was higher than children (264). The F0 interval between the means of L vs. M, M vs. H, and L vs. H were (in octaves), in order, 0.88, 0.54, and 1.42; cf. men (0.76), women (0.24), and children (1.00). These variations in pitch height across stimuli notwithstanding, all three stimuli (L, M, H) are representative of the same tonal category in Mandarin.

Iterated rippled noise (IRN) was used to create these stimuli by applying polynomial equations that generate dynamic, curvilinear pitch patterns (Swaminathan, Krishnan and Gandour, 2008). IRN enables us to preserve dynamic variations in pitch of auditory stimuli that lack formant structure, temporal envelope, and recognizable timbre characteristic of speech. IRN stimuli were created by delaying Gaussian noise (80–4000 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 (Patterson, Handel, Yost and Datta, 1996; Yost, 1996b; Krishnan, Bidelman and Gandour, 2010) with little or no change in salience beyond an iteration step of 32 (Yost, 1996a), the upper limit used here.

The experimental paradigm consisted of three segments: a 250 ms pitch segment (L, M, H) preceded by a 900 ms noise segment and followed by a 100 ms noise segment. Noise to pitch, and pitch to noise segment transitions were cross-faded using 7.5 ms cos² ramps to achieve the same overall root mean square value. This enabled us to avoid evoked potentials elicited by acoustic change between segments. The noise precursor segment enabled us to isolate the pitch specific response from the obligatory onset response (for details about paradigm, see Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner, 2003; Krishnan, Gandour, Ananthakrishnan and Vijayaraghavan, 2014). Similarly, the noise segment that follows the pitch segment serves to disentangle pitch offset response from sound offset response.

All stimuli were presented binaurally at 80 dB SPL using magnetically-shielded tubal insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA) with rarefaction onset polarity and a repetition rate of 0.56/s. Stimulus presentation order was randomized both within and across participants. Stimuli were generated and played out using an auditory evoked potential system (SmartEP, Intelligent Hearing Systems; Miami, FL, USA).

Cortical and brainstem evoked response data acquisition

Participants reclined comfortably in an electro-acoustically shielded booth. 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 of their choice (minus subtitles) throughout the recording session. The EEG was acquired continuously (5000 Hz sampling rate; 0.3 to 2500 Hz analog band-pass) through the ASA-Lab EEG 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. All other acquisition parameters are identical to those described in our previous publications (Krishnan, Gandour and Suresh, 2015b, a; 2016, details available upon request). For each stimulus, EEGs were acquired in two blocks of 750 sweeps each. The experimental protocol took about 2 hours to complete. All CPR and FFR analysis was performed offline using the *Experiment Manager* software (ASA, Version 9.2, ANT, The Netherlands).

Extraction of cortical and brainstem responses

CPR—EEG files were down sampled to 1024 Hz and 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 to extract the CPR. Sweeps containing electrical activity exceeding $\pm 50 \mu\text{V}$ were rejected automatically. Subsequently, averaging was performed on all 10 unipolar electrode locations (Fpz, AFz, Fz, F3, F4, Cz, T7, T8, M1, M2) using the common reference to allow comparison of CPR components at the right temporal (T8), and left temporal (T7) electrode sites to evaluate asymmetry effects.

In previous crosslanguage CPR studies, we have consistently observed robust differences in CPR neural activity over the T7 and T8 electrode sites that reflect a functional, experience-dependent rightward asymmetry (Krishnan, Gandour, Ananthakrishnan and Vijayaraghavan, 2014, 2015; Krishnan, Gandour and Suresh, 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 frontocentral sites (e.g., Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner, 2003; Bidelman and Grall, 2014; Krishnan, Gandour and Suresh, 2015b). It also allows us to compare our CPR data with Fz-derived POR data (Gutschalk, Patterson, Rupp, Uppenkamp and Scherg, 2002; Gutschalk, Patterson, Scherg, Uppenkamp and Rupp, 2004; Bidelman and Grall, 2014). While the analysis epoch was 1600 ms, including the 100 ms pre-stimulus baseline for both averages, only CPR responses to the pitch segment was analyzed.

FFR—EEG files, with the original 5000 Hz sampling frequency, were digitally band-pass filtered (75–1500 Hz, Butterworth zero phase filters with 24 dB/Octave rejection rate) followed by automatic rejection of sweeps containing electrical activity exceeding $\pm 40 \mu\text{V}$. To obtain the FFR, time domain averaging was then performed on three different re-referenced electrode montages (FPz-linked mastoids; Fz-linked mastoids; Cz-linked mastoids) over an analysis window of 270 ms from 893 (onset of pitch segment) to 1163 ms. Each FFR waveform represents the grand average of the FFRs derived from the three electrode montages using 1500 sweeps presented in two block of 750 sweeps each. This multichannel averaging served to improve detectability of the FFR by improving the signal to noise ratio.

CPR latency and magnitude

Peak latency 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 height on the CPR components. Individual averaged responses were overlaid on the grand averaged response to facilitate response detection and improve accuracy of manual peak picking for latency and amplitude measurements. The time instant proximal to the peak with maximum voltage was taken as the measure of absolute latency and peak amplitude for a given component. Interjudge reliability showed a high percentage of agreement (90%) in measurements made independently by two members of the laboratory. To evaluate response asymmetry at the temporal electrode sites (T7/T8), peak-to-peak amplitude of Na–Pb and Pb–Nb was measured for each stimulus condition. To visualize the asymmetry effects along a spectrotemporal dimension, a joint time frequency analysis of a continuous wavelet transform was performed on grand average waveforms derived from the temporal electrodes.

FFR neural pitch strength

The magnitude of the F0 component from each FFR waveform was used to quantify neural periodicity strength. To satisfy assumptions of normality and homogeneity of variance, raw F0 data was converted to \log_{10} scale. Fast Fourier transform of a windowed version of each FFR waveform (Gaussian window, 1 Hz resolution) was computed to extract the spectrum of the response. For each subject, the magnitude of the response peak at F0 was measured

relative to the noise floor. All FFR data analyses were performed using custom routines coded in Matlab 11 (The MathWorks, Inc., Natick, MA, USA).

Comparison of CPR and FFR

Because CPR components are several orders of magnitude larger than the FFR response, it is necessary to normalize the raw amplitude measures to allow for a meaningful comparison of the change in pitch-related neural activity at the brainstem and cortical levels independent of their differences in absolute magnitude. To measure the normalized, mean response magnitude of mid and low pitch height per component (Na–Pb, Pb–Nb, FFR), each participant's data values were divided by his/her maximum value pooling across pitch heights. This procedure enabled us to compare brainstem and cortical response components directly while preserving relative differences in magnitude between stimuli. The stimulus with H pitch height did not elicit discernible FFRs with good repeatability in all subjects, presumably reflecting degraded neural phase locking at the higher frequency. Comparison of the CPR and FFR was therefore restricted to stimuli with L and M pitch height.

Statistical analysis

At the Fz-linked T7/T8 electrode site, two-way (group x stimulus) mixed model ANOVAs (SAS®; SAS Institute, Inc., Cary, NC, USA) were performed on each component of peak latency (Na, Pb, Nb) and peak-to-peak amplitude (Na–Pb, Pb–Nb); at the temporal electrode sites (T7/T8), a two-way (group x electrode site) ANOVA was performed on peak-to-peak amplitude. Group (Chinese, English) served as the between-subjects factor; subjects nested within group served as the random factor. Stimulus (L, M, H) and electrode site (T7 [left], T8 [right]) were treated as within-subject factors. Post hoc multiple comparisons were adjusted with a Bonferroni significance level set at $\alpha = 0.05$. Partial eta-squared (η_p^2) values were reported to indicate effect size. Normality and homogeneity of variance assumptions were confirmed prior to statistical inference.

RESULTS

Response morphology of CPR components

Over Fz-linked T7/T8 sites, CPR components (Na, Pb, Nb) of the pitch segment are plainly distinguishable (Fig. 2, top). Across stimulus conditions (L, M, H), Na–Pb amplitude appears to be larger in Chinese than English listeners. CPR latency of Na, Pb, and Nb decrease with an increase in pitch height regardless of group.

Fz-linked T7/T8: Latency (Na, Pb, Nb) and amplitude (Na–Pb, Pb–Nb)

Mean peak latency of CPR components elicited by the pitch height stimuli (L, M, H) are displayed in Fig. 3 (top panel). For Na, ANOVA yielded a main effect of stimulus only ($F_{2,52} = 440.29$, $p < 0.0001$, $\eta_p^2 = 0.944$). Irrespective of group, post hoc multiple comparisons showed that peak latency was longer in response to the low pitch height compared to the mid (L vs M: $t_{52} = 16.20$, $p < 0.0001$) which, in turn, elicited a longer latency than high (M vs H: $t_{52} = 13.43$). This progressive decrease in latency from low to high entails that low pitch height is also longer than high (L vs H: $t_{52} = 29.63$). In the case of Pb and Nb, ANOVA

similarly yielded a stimulus main effect only (Pb: $F_{2,52} = 176.01$, $p < 0.0001$, $\eta_p^2 = 0.871$; Nb: $F_{2,52} = 29.58$, $p < 0.0001$, $\eta_p^2 = 0.532$). However, the pattern of contrasts among pitch height levels was different. Regardless of group, post hoc t tests indicated that both low and mid pitch height evoked longer response peak latency than high as reflected by both Pb (L vs H, $t_{52} = 17.11$, $p < 0.0001$; M vs H, $t_{52} = 15.22$) and Nb (L vs H, $t_{52} = 6.22$; M vs H, $t_{52} = 7.02$). For either component (Pb, Nb), latency differences between low and mid pitch height failed to reach significance.

Fig. 3 (middle panel) displays mean peak-to-peak amplitude of CPR components across pitch heights. As measured by Na–Pb, ANOVA revealed main effects of group ($F_{1,26} = 14.98$, $p = 0.0007$, $\eta_p^2 = 0.365$) and stimulus ($F_{2,52} = 13.44$, $p < 0.0001$, $\eta_p^2 = 0.341$). Post hoc comparison showed that amplitude was larger in Chinese than English listeners, pooled across pitch heights ($t_{26} = 3.87$, $p = 0.0007$). Regardless of group, we observed the same pattern of stimulus contrasts. Post hoc comparisons indicated that amplitude was larger in response to mid and high pitch heights relative to low (L vs M: $t_{52} = -5.03$, $p < 0.0001$; L vs H: $t_{52} = -3.59$, $p = 0.0022$). In the case of Pb–Nb (bottom panel), ANOVA yielded a main effect of stimulus only ($F_{2,52} = 16.10$, $p < 0.0001$, $\eta_p^2 = 0.382$). Like Na–Pb, mid and high pitch heights similarly evoked larger amplitude than low (L vs M: $t_{52} = -4.06$, $p = 0.0005$; L vs H: $t_{52} = -5.46$, $p < 0.0001$). With respect to Fz amplitude, these findings collectively reveal that language-dependent effects on sensitivity to changes in pitch height are restricted to Na–Pb.

T7/T8: Amplitude (Na–Pb, Pb–Nb)

Fig. 4 displays grand average waveforms of CPR components (left) evoked by pitch height stimuli (L, M, H) for each language group and their matching spectra (right). Chinese listeners show a rightward asymmetry (T8 > T7) as seen in both waveforms and spectrotemporal plots across pitch height levels.

Fig. 5 shows Na–Pb (top) and Pb–Nb (bottom) peak-to-peak amplitude—pooled across pitch heights—per language group and temporal electrode site (T7, T8). For Na–Pb, the omnibus two-way ANOVA revealed a main effect of electrode site ($F_{1,26} = 8.31$, $p = 0.0078$, $\eta_p^2 = 0.242$) and group x electrode site interaction ($F_{1,26} = 9.50$, $p = 0.0048$, $\eta_p^2 = 0.268$). By group, a rightward asymmetry was observed in the Chinese group only ($t_{26} = -4.22$, $p = 0.0003$). By electrode site, amplitude of the Chinese was larger than the English group over the right electrode site only ($t_{26} = 3.00$, $p = 0.0060$). For Pb–Nb, the omnibus ANOVA yielded a group x electrode site interaction ($F_{1,26} = 7.00$, $p = 0.0136$, $\eta_p^2 = 0.212$). Similar to Na–Pb, simple effects of group revealed a rightward asymmetry in the Chinese group only ($t_{26} = -3.28$, $p = 0.0029$). These findings together demonstrate that language experience effects on sensitivity to changes in pitch height are evident in both CPR components, but restricted to pitch-relevant neural activity over T8, the right temporal electrode site.

Temporal and spectral response characteristics of FFR

Fig. 6 displays grand-averaged FFR waveforms (left) elicited by pitch height stimuli (L, M) are shown for each language group and their matching spectra (right). In both groups, the waveform and F0 spectral magnitude are larger for the low pitch height stimulus compared to the mid. However, response magnitude of the Chinese group appears to be relatively more robust irrespective of stimulus.

Neural pitch strength of FFR as a function of pitch height

Fig. 7 shows mean F0 magnitude FFR encoding at each of two steps (L, M) along the pitch height continuum. An omnibus ANOVA performed on normalized (\log_{10}) data revealed significant main effects of group ($F_{1,26} = 15.28, p = 0.0006, \eta_p^2 = 0.370$) and stimulus ($F_{1,26} = 97.79, p < 0.0001, \eta_p^2 = 0.790$). The group x stimulus interaction was not significant ($F_{1,26} = 0.87, p = 0.3591$). These results mean that F0 magnitude was larger in Chinese than in English listeners regardless of pitch height; low pitch height elicited greater F0 magnitude than high across groups.

Comparison of CPR and FFR responses

Fig. 8 displays mean normalized response amplitude of each group per component and stimulus. An omnibus three-way (group x component x stimulus) ANOVA on normalized amplitudes showed a strong interaction between component and stimulus ($F_{2,52} = 76.71, p < 0.0001, \eta_p^2 = 0.747$). By CPR components, mid pitch height was greater than low (Na–Pb: $t_{52} = -5.98, p < 0.0001$; Pb–Nb: $t_{52} = -6.50, p < 0.0001$). By the brainstem FFR component, just the reverse was true; low pitch height was greater than mid ($t_{52} = 8.06, p < 0.0001$). By stimulus, low pitch height evoked larger amplitude in Na–Pb ($t_{52} = 3.29, p = 0.0055$) and FFR ($t_{52} = -2.66, p = 0.0309$) relative to Pb–Nb. The mid pitch height, on the other hand, evoked larger amplitude in Na–Pb compared to either Pb–Nb ($t_{52} = 2.88, p = 0.0173$) or FFR ($t_{52} = 11.58, p < 0.0001$). Moreover, mid pitch height evoked larger amplitude in Pb–Nb than FFR ($t_{52} = 8.70, p < 0.0001$). These data together suggest that the response amplitude differs between FFR and CPR components as a function of increasing pitch height. This contrast between FFR and CPR likely means that two different mechanisms underlie responses at brainstem and cortical levels. FFR reflects sustained neural phase locking; CPR, on the other hand, reflects transient neural activity synchronized to certain temporal attributes of pitch.

DISCUSSION

These findings demonstrate that language universal and language dependent factors influence neural representation of information relevant to pitch height. Regardless of language experience, CPR latency and amplitude show an essentially identical pattern with increasing pitch height: i.e., a decrease in latency (Na, Pb, Nb); an increase in amplitude (Na–Pb, Pb–Nb). Language dependent (C > E) changes with pitch height are observed only for Na–Pb at both Fz and Fz-linked T7/T8 sites. The temporal electrode sites also show a rightward asymmetry (T8 > T7) for native Chinese. In the brainstem, language dependent (C > E) changes in FFR F0 magnitude decrease with increasing pitch height. CPR and FFR

responses reveal opposite patterns of relative changes in magnitude across pitch heights. FFR decreases in magnitude; CPR increases. Differences in the pattern of changes in magnitude suggest that two different pitch encoding schemes operate in the brainstem and auditory cortex.

Language-universal sensitivity of the CPR and FFR to changes in pitch height

Our findings show that irrespective of language group, latency of CPR components, especially Na, decreases with increasing pitch height. This shortening effect is consistent with previous reports of a shortening of cortical response latency elicited by increasing pitch height: cortical N100 component, transition from silence to sound (Ragot and Lepaul-Ercole, 1996; Roberts and Poeppel, 1996); POR, transition from noise to pitch (Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner, 2003; Ritter, Gunter Dosch, Specht and Rupp, 2005; Ritter, Dosch, Specht, Schneider and Rupp, 2007; Bidelman, 2015). Systematic decrease in response latency with increasing pitch height reflects, in part, the traveling wave delay along the cochlear partition and tonotopic organization in the auditory cortex (Pantev, Hoke, Lutkenhoner and Lehnertz, 1989; Pantev, Bertrand, Eulitz, Verkindt, Hampson, Schuierer and Elbert, 1995; Bidelman and Grall, 2014). An alternative view is that the tonotopic map in primary auditory cortex is actually a periodotopic map (Pantev, Hoke, Lutkenhoner and Lehnertz, 1989; Pantev, Bertrand, Eulitz, Verkindt, Hampson, Schuierer and Elbert, 1995). But it remains unclear how a periodotopic map may contribute to pitch processing (cf. Langner, Sams, Heil and Schulze, 1997). It is incontrovertible that cochlear delay and the tonotopic arrangement in primary auditory cortex contribute to latency shifts with pitch height. Nevertheless, it is also clear that they do not completely account for the relatively longer latency of the POR as compared to the response elicited to onset of sounds in silence.

This longer latency for CPRs suggest that pitch-specific computations are integrated over several cycles of the regular interval stimulus. Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner (2003) have proposed a function relating POR response latency and pitch height. In this function, the auditory system has to integrate over a duration of at least four times the delay to estimate pitch of IRN stimuli. Changes in latency with pitch height in this study are in line with this prediction. Consistent with the notion that the POR latency changes do not simply represent place of cochlear activity alone, Ritter et al. (2005) conclude that the latency differences in the cortical N100 for positive gain and negative gain IRN stimuli represents a neurophysiologic correlate of perceived pitch. Finally, our results fail to show significant differences in latency between low and mid pitch height stimuli as reflected by Pb and Nb. Such findings are congruent with the notion that CPR components are differentially sensitive to specific temporal pitch attributes (e.g., pitch salience; Krishnan, Gandour and Suresh, 2016). Overall, these results suggest that temporal integration mechanisms to extract pitch information is fundamentally similar in both language groups.

With respect to response magnitude, the pattern of changes with increasing pitch height in the cortical CPR components and the brainstem FFR were similar in both groups. Na–Pb and Pb–Nb amplitude tends to increase with pitch height; high and mid has relatively greater

amplitude than low. Conversely, F0 magnitude of the FFR decreases with increasing pitch height. Our CPR data are mostly consistent with previous studies of pitch height effects on the magnitude of pitch-relevant neural activity in auditory cortex. Using IRN stimuli with either 2 or 8 iteration steps, POR amplitude increases with increasing pitch height (Ritter, Gunter Dosch, Specht and Rupp, 2005). POR amplitude also increases with a delay of IRN stimuli from 16 ms (64 Hz) to 8 ms (125 Hz); shorter delays are ineffective (Krumbholz, Patterson, Seither-Preisler, Lammertmann and Lutkenhoner, 2003). Bidelman (2015), to the contrary, reported no difference in POR magnitude elicited by IRN stimuli with a low and high pitch height. This lack of a pitch height effect is attributable to large inter- and intra-subject variability in their POR amplitude data. It appears that the increase in CPR amplitude with increasing pitch height most likely represents an increase in temporal synchronization of pitch-relevant neural activity as the temporal window for pitch extraction decreases. Our finding that FFR F0 magnitude decreases with increasing pitch height is consistent with earlier reports (Batra, Suwada and Maer, 1986; Hoormann, Falkenstein, Hohnsbein and Blanke, 1992; Krishnan, 2002), and suggests that neural phase-locking ability decreases with increasing frequency (Rose, Brugge, Anderson and Hind, 1967; Lavine, 1971). Collectively, these findings suggest that two distinct, language-universal sensory level neural mechanisms, respectively, underlie pitch-relevant processing at the auditory brainstem (temporal code based on neural phase locking) and auditory cortex (discharge rate based code).

Language-dependent enhancement of cortical and brainstem representation of pitch height

We observed a robust Chinese group response enhancement for both the brainstem (FFR) and cortical CPR component Na–Pb across stimuli varying in pitch height. Such enhancement likely reflects a sharpening of response properties of neural populations to ensure optimal representation of behaviorally-relevant pitch attributes. These data further reinforce our earlier observations of language experience-dependent shaping of pitch representation at both cortical and brainstem levels of the brain (for reviews, see Krishnan, Gandour and Bidelman, 2012; Gandour and Krishnan, 2014; Krishnan and Gandour, 2014; Gandour and Krishnan, 2016).

Changes in pitch height elicit responses that are experience-dependent or experience-independent. In our view, the Chinese group's enhancement of pitch-relevant neural activity represents modulation of early sensory pitch mechanisms by extrasensory processes at a higher hierarchical level in the brainstem and auditory cortex (for additional discussion, see Krishnan, Gandour and Suresh, 2016, p. 111). These adaptive, cortical pitch mechanisms sharpen neural response properties to optimize representation of temporal attributes of pitch. According to a theoretical framework for experience-induced neural plasticity (Krishnan, Gandour and Bidelman, 2012), the experience-dependent effect at the cortical level may reflect, in part, enhanced fine-grained output from brainstem pitch mechanisms. This coordination between auditory brainstem and cortex has also been reported in studies of long-term language experience with French syllables (Intartaglia, White-Schwoch, Meunier, Roman, Kraus and Schon, 2016); and older musicians' experience with English vowel perception (Bidelman and Alain, 2015).

Why are language-dependent responses to changes in pitch height restricted to Na–Pb only? Whenever pitch acceleration or the shape of pitch contours is manipulated, experience-dependent enhancement applies to both Na–Pb and Pb–Nb (Krishnan, Gandour, Ananthakrishnan and Vijayaraghavan, 2015; Krishnan, Gandour and Suresh, 2015b, a). If, on the other hand, changes in acceleration or shape are held constant, language-dependent effects are restricted to the earlier Na–Pb time window only (e.g., pitch salience; Krishnan, Gandour and Suresh, 2016). Herein, we manipulated only pitch height with no change in pitch acceleration or shape. Like pitch salience, experience-dependent enhancement of pitch height is restricted to Na–Pb. It is plausible that Na–Pb optimally represents neural processing relevant to extraction of an estimate of pitch and its strength (e.g., pitch salience, pitch height). The Pb–Nb time window, however, may index other dynamic, time-variant pitch attributes. Thus, selective enhancement of specific CPR components may suggest utilization of temporal integration windows optimized for processing a specific attribute of pitch. Taken together, pitch processing along the auditory neuraxis involves a coordinated interplay between sensory and extrasensory processes. Their relative weighting are determined by both language experience and neural sensitivity to particular pitch attributes within a given temporal integration window.

A rightward asymmetry over the temporal electrode sites is restricted to Chinese listeners. We observe this asymmetry in Na–Pb amplitude, but not Pb–Nb. Moreover, Chinese Na–Pb amplitude is larger than that of English over the right temporal site only. These results are consistent with the well-established body of empirical evidence supporting right hemisphere preference in processing pitch, linguistic or otherwise (Meyer, 2008; Zatorre and Gandour, 2008; Friederici, 2011). The experience-dependent rightward asymmetry shows that extrasensory components may mask sensory effects within a specific time window.

A simple invocation of more fine-grained processing for pitch in the right hemisphere would predict relative rightward asymmetry regardless of language experience. Using the CPR, we have found that language-dependent effects apply only to pitch contours that exhibit temporal attributes characteristic of natural speech. We infer that such effects result from experience-driven, extrasensory modulation of adaptive pitch mechanisms at early sensory levels of pitch processing in the right auditory cortex. In this case, these adaptive mechanisms in the right hemisphere afford more fine-grained processing of linguistically relevant, temporal pitch attributes. Such a formulation is consistent with our recent observations of rightward asymmetry at temporal electrode sites (Krishnan, Gandour, Ananthakrishnan and Vijayaraghavan, 2014, time-variant vs. time-invariant; Krishnan, Gandour and Suresh, 2015b, rate of pitch acceleration; pitch direction & location of peak acceleration, Krishnan, Gandour and Suresh, 2015a; 2016, salience).

Different pitch encoding schemes in the auditory brainstem and cortex

Pitch-relevant neural activity elicited by low and mid pitch height reveals different patterns of change in magnitude between the auditory brainstem and cortex (Fig. 8). CPR reflects transient neural activity synchronized to certain temporal attributes of pitch; in contrast, FFR reflects sustained neural phase locking. Growth in amplitude of CPR components likely reflects a frequency dependent increase in neural synchrony that increases the firing rate.

These data are compatible with the operation of two distinct neural schemes representing pitch. In the brainstem, periodicity and pitch are instantiated by sustained neural phase locking to the fundamental periodicity of the stimulus waveform (Cariani and Delgutte, 1996a, b; Plack, 2005). Neurons in primary auditory cortex exhibit temporal and spectral response properties capable of implementing both pitch-encoding schemes (Steinschneider, Reser, Fishman, Schroeder and Arezzo, 1998; Lu, Liang and Wang, 2001). However, the question remains whether or not they form a network with neurons that are selective to pitch. Within cortical neurons, it is more common to see modulation of spike rates vary as a function of F0 in the encoding of periodicity and pitch. Walker, Bizley, King and Schnupp (2011) hypothesize that cortical neurons may be too sluggish to provide phase-locked representations of periodicity within the pitch range because of wider temporal integration windows in auditory cortex. It therefore remains an empirical question as to how pitch-relevant information is extracted at these two levels of biological structure. Specifically, how is pitch information transformed from a temporal phase-locking code in the brainstem to a spike rate code in cortical neurons?

Conclusions

By systematically varying pitch height, we are able to distinguish pitch-relevant neural activity attributable to language-independent sensory processes from overlaid language-dependent neural activity. A similar pattern of changes in CPR latency/amplitude and FFR magnitude point to shared, neural mechanisms underlying pitch encoding independent of language experience. A language-dependent effect is demonstrated by enhanced sensitivity to pitch height in the Chinese group regardless of level of brain structure, and a rightward asymmetry over temporal electrode sites as reflected by Na–Pb. The selectivity of Na–Pb suggests that extrasensory modulation of sensory processes are targeted to response components and/or temporal windows that best index a specific temporal attribute of pitch. Differential sensitivity to pitch height between CPR and FFR responses may imply a transformation from a temporal pitch encoding scheme in the brainstem to a discharge rate pitch encoding scheme at the cortical level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This research was supported by the National Institutes of Health 5R01DC008549-08 (A.K.). Thanks to Rongrong Zhang and Jincheng Bai for help with statistical analysis; Whitney Lyle, for assistance with data acquisition.

The authors declare no conflict of interest.

Author contributions: A.K., C.H.S., and J.T.G. designed research and analyzed data; A.K. and C.H.S. performed research; A.K. and J.T.G. wrote the paper.

List of Abbreviations

ANOVA analysis of variance

C	Chinese
CPR	cortical pitch-specific response
E	English
EEG	electroencephalography
FFR	frequency following response
F0	voice fundamental frequency
H	high pitch height
IRN	iterated rippled noise
L	low pitch height
M	mid pitch height
MMN	mismatch negativity
MEG	magnetoencephalography
POR	pitch onset response

References

- Batra R, Suwada S, Maer VL. Human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. *Hear Res.* 1986; 59:179–188.
- Bidelman GM. Sensitivity of the cortical pitch onset response to height, time-variance, and directionality of dynamic pitch. *Neurosci Lett.* 2015; 603:89–93. DOI: 10.1016/j.neulet.2015.07.018 [PubMed: 26200250]
- Bidelman GM, Alain C. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J Neurosci.* 2015; 35:1240–1249. DOI: 10.1523/JNEUROSCI.3292-14.2015 [PubMed: 25609638]
- Bidelman GM, Grall J. Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *Neuroimage.* 2014; 101:204–214. DOI: 10.1016/j.neuroimage.2014.07.005 [PubMed: 25019679]
- Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol.* 1996a; 76:1698–1716. [PubMed: 8890286]
- Cariani PA, Delgutte B. 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.* 1996b; 76:1717–1734. [PubMed: 8890287]
- Chandrasekaran B, Gandour JT, Krishnan A. Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. *Restor Neurol Neurosci.* 2007; 25:195–210. [PubMed: 17942999]
- Francis AL, Ciocca V, Ma L, Fenn K. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *J Phonetics.* 2008; 36:268–294. DOI: 10.1016/j.wocn.2007.06.005
- Friederici AD. The brain basis of language processing: from structure to function. *Physiol Rev.* 2011; 91:1357–1392. DOI: 10.1152/physrev.00006.2011 [PubMed: 22013214]
- Gandour JT. Tone perception in Far Eastern languages. *J Phonetics.* 1983; 11:149–175.
- Gandour, JT. Phonetics of tone. In: Asher, R., Simpson, J., editors. *The encyclopedia of language & linguistics.* Pergamon Press; New York: 1994. p. 3116-3123.

- Gandour, JT., Krishnan, A. Neural bases of lexical tone. In: Winskel, H., Padakannaya, P., editors. *Handbook of South and Southeast Asian psycholinguistics*. Cambridge University Press; Cambridge, UK: 2014. p. 339-349.
- Gandour, JT., Krishnan, A. Processing tone languages. In: Hickok, G., Small, SL., editors. *Neurobiology of language*. Academic Press; New York: 2016. p. 1095-1107.
- Griffiths TD, Warren JD. The planum temporale as a computational hub. *Trends Neurosci*. 2002; 25:348–353. [PubMed: 12079762]
- Gutschalk A, Patterson RD, Rupp A, Uppenkamp S, Scherg M. Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage*. 2002; 15:207–216. DOI: 10.1006/nimg.2001.0949 [PubMed: 11771990]
- Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A. Temporal dynamics of pitch in human auditory cortex. *Neuroimage*. 2004; 22:755–766. S1053811904000680 [pii]. DOI: 10.1016/j.neuroimage.2004.01.025 [PubMed: 15193604]
- Hoormann J, Falkenstein M, Hohnsbein J, Blanke L. The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. *Hear Res*. 1992; 59:179–188. [PubMed: 1618709]
- Huang T, Johnson K. Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*. 2011; 67:243–267.
- Intartaglia B, White-Schwoch T, Meunier C, Roman S, Kraus N, Schon D. Native language shapes automatic neural processing of speech. *Neuropsychologia*. 2016; 89:57–65. DOI: 10.1016/j.neuropsychologia.2016.05.033 [PubMed: 27263123]
- Krishnan A. Human frequency-following responses: representation of steady-state synthetic vowels. *Hear Res*. 2002; 166:192–201. [PubMed: 12062771]
- Krishnan A, Bidelman GM, Gandour JT. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hear Res*. 2010; 268:60–66. DOI: 10.1016/j.heares.2010.04.016 [PubMed: 20457239]
- Krishnan A, Bidelman GM, Smalt CJ, Ananthakrishnan S, Gandour JT. Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*. 2012; 50:2849–2859. DOI: 10.1016/j.neuropsychologia.2012.08.013 [PubMed: 22940428]
- Krishnan A, Gandour JT. Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: electrophysiological evidence. *Acoustics Australia*. 2014; 42:166–178. [PubMed: 25838636]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V. Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*. 2014; 59:1–12. DOI: 10.1016/j.neuropsychologia.2014.04.006 [PubMed: 24751993]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V. Language experience enhances early cortical pitch-dependent responses. *J Neurolinguistics*. 2015; 33:128–148. DOI: 10.1016/j.jneuroling.2014.08.002 [PubMed: 25506127]
- Krishnan A, Gandour JT, Bidelman GM. Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport*. 2012; 23:498–502. DOI: 10.1097/WNR.0b013e328353764d [PubMed: 22495037]
- Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*. 2009; 20:408–413. DOI: 10.1097/WNR.0b013e3283263000 [PubMed: 19223793]
- Krishnan A, Gandour JT, Suresh CH. Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. *Brain Lang*. 2014; 138:51–60. <http://dx.doi.org/10.1016/j.bandl.2014.09.005>. [PubMed: 25306506]
- Krishnan A, Gandour JT, Suresh CH. Experience-dependent enhancement of pitch-specific responses in the auditory cortex is limited to acceleration rates in normal voice range. *Neuroscience*. 2015a; 303:433–445. DOI: 10.1016/j.neuroscience.2015.07.015 [PubMed: 26166727]
- Krishnan A, Gandour JT, Suresh CH. Pitch processing of dynamic lexical tones in the auditory cortex is influenced by sensory and extrasensory processes. *Eur J Neurosci*. 2015b; 41:1496–1504. DOI: 10.1111/ejn.12903 [PubMed: 25943576]

- Krishnan A, Gandour JT, Suresh CH. Language-experience plasticity in neural representation of changes in pitch salience. *Brain Res.* 2016; 1637:102–117. DOI: 10.1016/j.brainres.2016.02.021 [PubMed: 26903418]
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lutkenhoner B. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb Cortex.* 2003; 13:765–772. [PubMed: 12816892]
- Krumhansl, CL. *Cognitive foundations of musical pitch.* Oxford University Press; New York: 1990. p. 112-114.
- Langner G, Sams M, Heil P, Schulze H. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *J Comp Physiol A.* 1997; 181:665–676. [PubMed: 9449825]
- Laver, J. *Principles of phonetics.* Cambridge University Press; New York, NY: 1994.
- Lavine RA. Phase-locking in response of single neurons in cochlear nuclear complex of the cat to low-frequency tonal stimuli. *J Neurophysiol.* 1971; 34:467–483. [PubMed: 5560042]
- Lu T, Liang L, Wang X. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci.* 2001; 4:1131–1138. DOI: 10.1038/nn737 [PubMed: 11593234]
- Melara RD, Marks LE. Processes underlying dimensional interactions: correspondences between linguistic and nonlinguistic dimensions. *Mem Cognit.* 1990; 18:477–495.
- Meyer M. Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitschrift für Neuropsychologie.* 2008; 19:101–115.
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia.* 1971; 9:97–113. [PubMed: 5146491]
- Pantev C, Bertrand O, Eulitz C, Verkindt C, Hampson S, Schuierer G, Elbert T. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalogr Clin Neurophysiol.* 1995; 94:26–40. [PubMed: 7530637]
- Pantev C, Hoke M, Lutkenhoner B, Lehnertz K. Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science.* 1989; 246:486–488. [PubMed: 2814476]
- Patterson RD, Handel S, Yost WA, Datta AJ. The relative strength of the tone and noise components in iterated ripple noise. *J Acoust Soc Am.* 1996; 100:3286–3294.
- Peterson GE, Barney HL. Control methods used in a study of the vowels. *J Acoust Soc Am.* 1952; 24:175–184.
- Plack, CJ. *Pitch: Neural coding and Pitch perception.* Springer; New York: 2005.
- Poeppl D, Roberts T. Effects of vowel pitch and task demands on latency and amplitude of the auditory evoked M100. *Brain Cogn.* 1996; 32:156–158.
- Ragot R, Lepaul-Ercole R. Brain potentials as objective indexes of auditory pitch extraction from harmonics. *Neuroreport.* 1996; 7:905–909. [PubMed: 8724670]
- Ritter S, Dosch HG, Specht HJ, Schneider P, Rupp A. Latency effect of the pitch response due to variations of frequency and spectral envelope. *Clin Neurophysiol.* 2007; 118:2276–2281. DOI: 10.1016/j.clinph.2007.06.017 [PubMed: 17709289]
- Ritter S, Gunter Dosch H, Specht HJ, Rupp A. Neuromagnetic responses reflect the temporal pitch change of regular interval sounds. *Neuroimage.* 2005; 27:533–543. DOI: 10.1016/j.neuroimage.2005.05.003 [PubMed: 15964207]
- Roberts TP, Poeppl D. Latency of auditory evoked M100 as a function of tone frequency. *Neuroreport.* 1996; 7:1138–1140. [PubMed: 8817518]
- Roffler SK, Butler RA. Localization of tonal stimuli in the vertical plane. *J Acoust Soc Am.* 1968; 43:1260–1266. [PubMed: 5659494]
- Rose JE, Brugge JF, Anderson DJ, Hind JE. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol.* 1967; 30:769–793. [PubMed: 4962851]

- Seither-Preisler A, Krumbholz K, Lutkenhoner B. Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? *Audiol Neurootol*. 2003; 8:322–337. DOI: 10.1159/000073517 [PubMed: 14566103]
- Shepard, RN. Structural representation of musical pitch. In: Deutsch, D., editor. *The psychology of music*. Academic Press; New York: 1982. p. 344-390.
- Somers M, Aukes MF, Ophoff RA, Boks MP, Fler W, de Visser KC, Kahn RS, Sommer IE. On the relationship between degree of hand-preference and degree of language lateralization. *Brain Lang*. 2015; 144:10–15. DOI: 10.1016/j.bandl.2015.03.006 [PubMed: 25880901]
- Steinschneider M, Reser DH, Fishman YI, Schroeder CE, Arezzo JC. Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *J Acoust Soc Am*. 1998; 104:2935–2955. [PubMed: 9821339]
- Swaminathan J, Krishnan A, Gandour JT. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*. 2008; 19:1163–1167. 00001756-200807160-00017 [pii]. DOI: 10.1097/WNR.0b013e3283088d31 [PubMed: 18596621]
- Tsang YK, Jia S, Huang J, Chen HC. ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neurosci Lett*. 2011; 487:268–272. S0304-3940(10)01374-1 [pii]. DOI: 10.1016/j.neulet.2010.10.035 [PubMed: 20970477]
- Walker KM, Bizley JK, King AJ, Schnupp JW. Cortical encoding of pitch: recent results and open questions. *Hear Res*. 2011; 271:74–87. DOI: 10.1016/j.heares.2010.04.015 [PubMed: 20457240]
- Wang XD, Wang M, Chen L. Hemispheric lateralization for early auditory processing of lexical tones: Dependence on pitch level and pitch contour. *Neuropsychologia*. 2013; 51:2238–2244. DOI: 10.1016/j.neuropsychologia.2013.07.015 [PubMed: 23911775]
- Warren JD, Uppenkamp S, Patterson RD, Griffiths TD. Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences*. 2003; 100:10038–10042. 1730682100 [pii]. DOI: 10.1073/pnas.1730682100
- Wong PC, Perrachione TK. Learning pitch patterns in lexical identification by native English-speaking adults. *Appl Psycholinguist*. 2007; 28:565–585.
- Xu Y. Contextual tonal variations in Mandarin. *J Phonetics*. 1997; 25:61–83.
- Xu Y, Krishnan A, Gandour JT. Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport*. 2006; 17:1601–1605. 00001756-200610230-00008 [pii]. DOI: 10.1097/01.wnr.0000236865.31705.3a [PubMed: 17001276]
- Yost WA. Pitch of iterated rippled noise. *J Acoust Soc Am*. 1996a; 100:511–518. [PubMed: 8675844]
- Yost WA. Pitch strength of iterated rippled noise. *J Acoust Soc Am*. 1996b; 100:3329–3335. [PubMed: 8914314]
- Zatorre RJ, Gandour JT. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos Trans R Soc Lond B Biol Sci*. 2008; 363:1087–1104. J412P80575385013 [pii]. DOI: 10.1098/rstb.2007.2161 [PubMed: 17890188]

- Cortical pitch-specific response (CPR) components (Na–Pb) index specific features of dynamic pitch (pitch height)
- Language-dependent weighting is sensitive to specific attributes of pitch within a given temporal integration window
- CPR amplitude increases with increasing pitch height; brainstem F0 magnitude decreases
- Similar pattern of changes from brainstem to cortex point to shared neural mechanisms regardless of language experience
- Differences between CPR and F0 magnitude imply distinct neural mechanisms for pitch extraction along auditory pathway

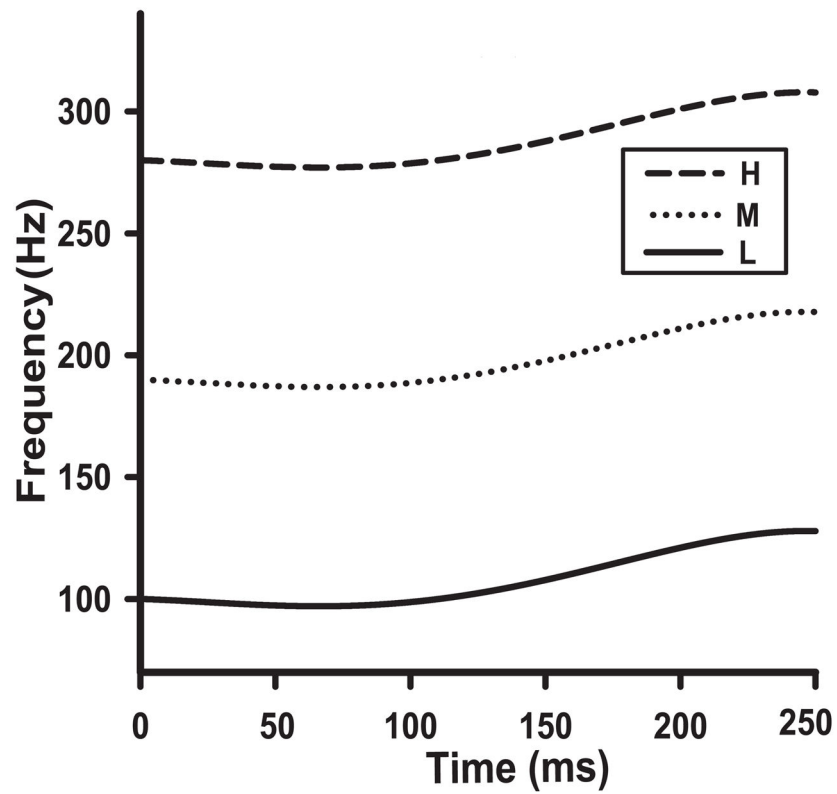


Figure 1. IRN stimuli used to evoke cortical responses to a pitch contour exemplary of Mandarin Tone 2. Voice fundamental frequency (F0) trajectories are displayed for three stimuli differentiated by varying degrees of height: low (L), mid (M), and high (H). They represent the pitch range for men, women, and children, respectively.

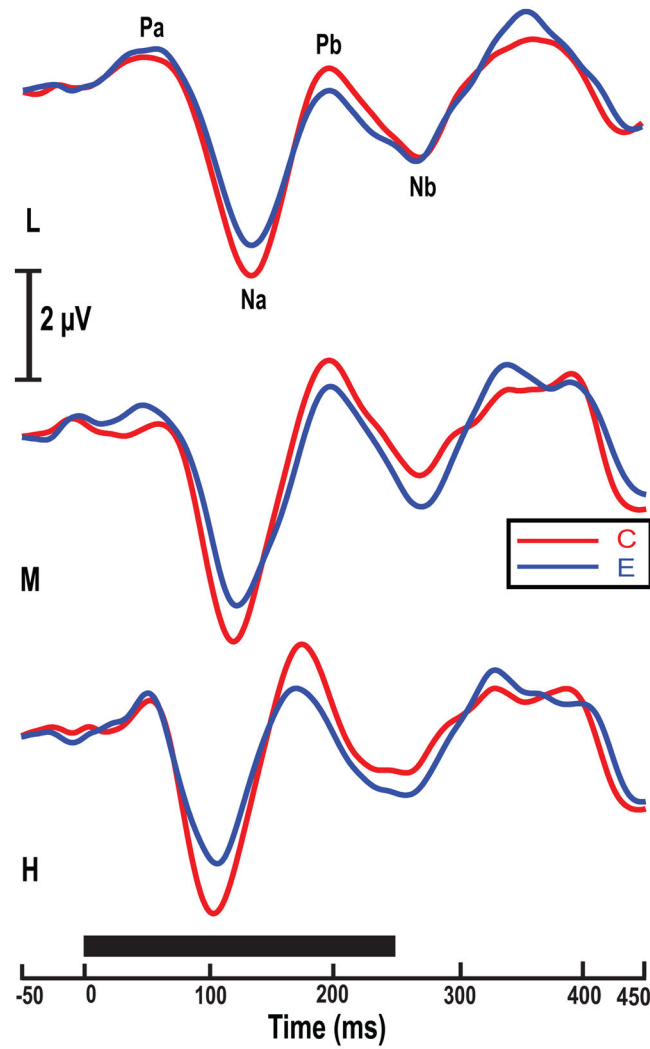


Figure 2. Grand average waveforms of the Chinese (C) and English (E) groups at the Fz-linked T7/T8 electrode site per pitch height level. Grand average waveforms of the Chinese and English groups at the Fz electrode site per stimulus condition (L, M, H). Na, Pb and Nb (highlighted in gray; top panel) are the most robust pitch-relevant components. CPR waveforms elicited by the three stimuli show that amplitude of Na, Pb and Nb appear to be more robust for the Chinese group across stimuli. Solid black horizontal bar indicates the duration of each stimulus.

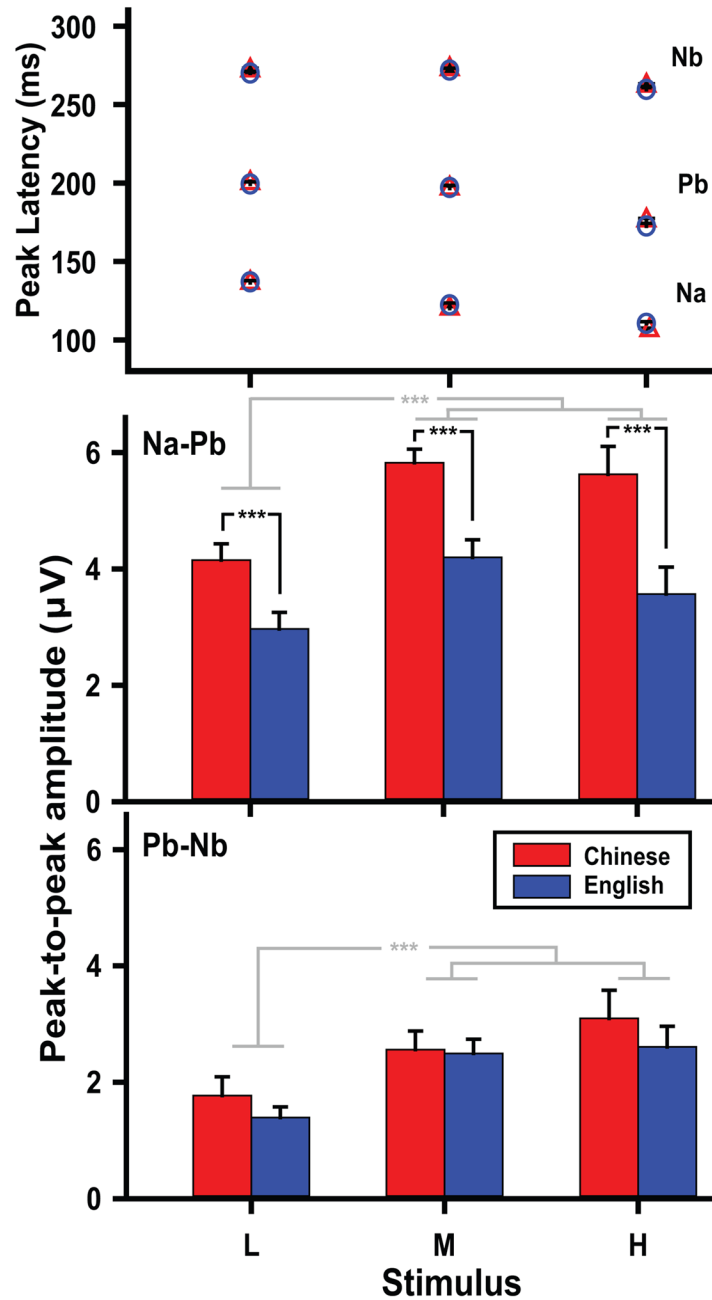


Figure 3. Mean peak latency of CPR components (Na, Pb, Nb) elicited by each pitch height level at the Fz-linked T7/T8 electrode site (top). In both Chinese (red triangle) and English (blue circle) groups, Na response latency is longer for low pitch height as compared to mid which, in turn, is longer than high. In contrast, Pb and Nb latency is longer for low and mid pitch height as compared to high. Language experience notwithstanding, peak latency gets shorter as the pitch gets higher. Mean peak-to-peak amplitude of CPR components show that Na–Pb amplitude (middle) is larger for Chinese listeners relative to English across pitch height levels (***) $p = 0.0007$). Regardless of group, mid and high, pitch heights are larger in

amplitude than low (***) $p < 0.0001$). Contrasts in pitch height for Pb–Nb (bottom) are identical to those for Na–Pb (***) $p < 0.0001$). CPR, cortical pitch response. Error bars = ± 1 SE.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

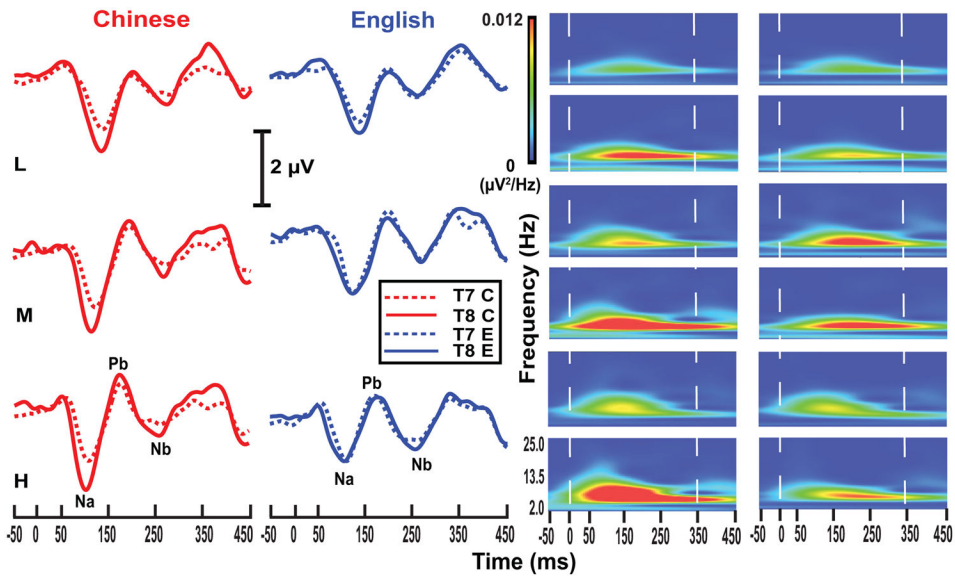


Figure 4. Grand average waveforms (left) and their corresponding spectra (right) of the CPR components for the two language groups (Chinese, red; English, blue) recorded at electrode sites T7 (dashed) and T8 (solid) for each of the three stimuli (L, M, H). CPR waveforms show a rightward asymmetry in favor of the right temporal electrode site for the Chinese group across pitch heights. The robust rightward asymmetry is evident in the spectrotemporal plots. 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 three stimuli.

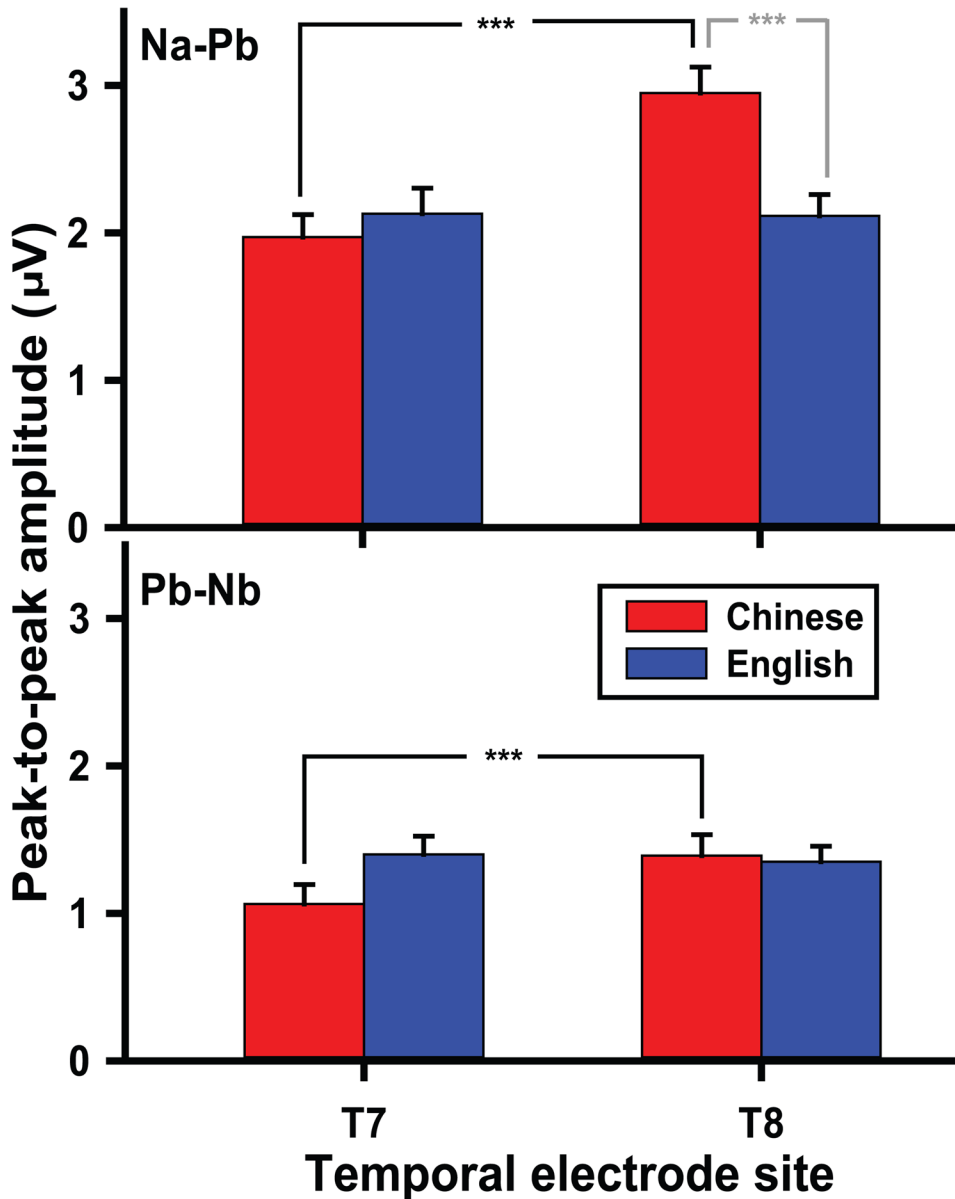


Figure 5.

Mean peak-to-peak amplitude of CPR components (Na–Pb, Pb–Nb) at the temporal electrode sites (T7, T8) in both Chinese and English groups. For Na–Pb (top), a rightward asymmetry (T8 > T7) is observed in the Chinese group only (***) $p < 0.0001$). At the right temporal site (T8), amplitude of the Chinese group is larger than that of the English (* $p = 0.0009$). For Pb–Nb (bottom), there are no remarkable differences in amplitude between language groups. Thus, the effects of language experience on sensitivity to changes in pitch height are limited to the Na–Pb component at the right temporal electrode site. Like Na–Pb, a rightward asymmetry appears in the Chinese group (***) $p < 0.0001$).

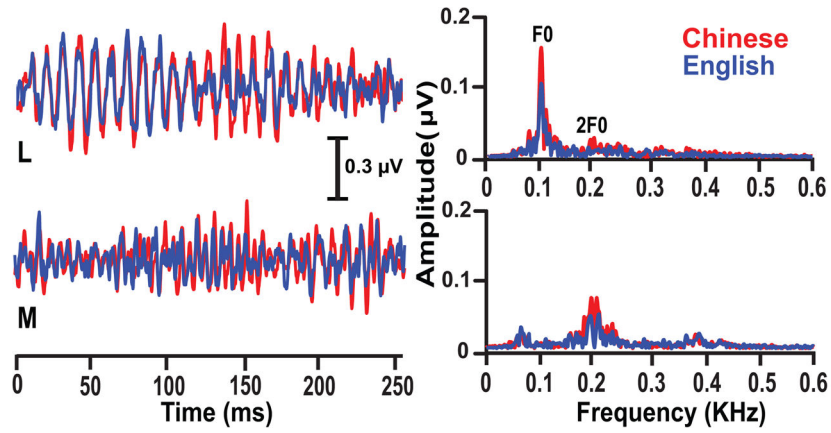


Figure 6. Grand average FFR waveforms (left) and their corresponding spectra (right) in response to the low (L) and mid (M) pitch height stimulus per language group. For both groups, the waveform and the F_0 spectral magnitude are larger for the L pitch height compared to the M. Responses, nonetheless, are relatively more robust for the Chinese group across pitch heights. The broader and smaller F_0 component for the M pitch height is consistent with the frequency-dependent decrease in neural phase-locking ability.

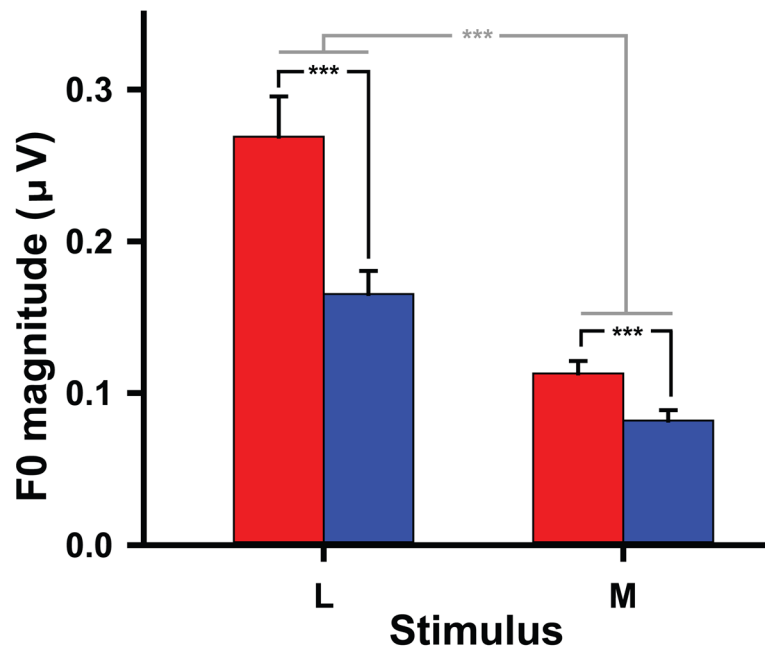


Figure 7. Mean F0 magnitude of the FFR for the Chinese (red) and English (blue) groups per L and M pitch height. F0 magnitude is larger in the Chinese than English group regardless of pitch height (** $p = 0.0006$). L elicited greater F0 magnitude than M across groups (** $p < 0.0001$).

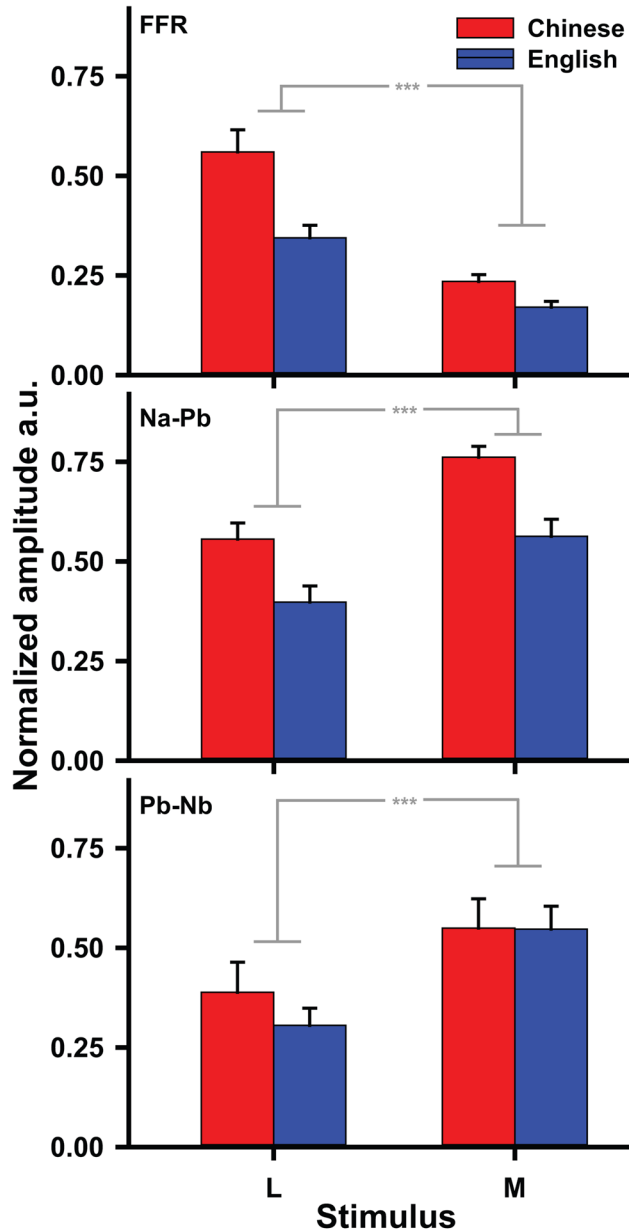


Figure 8. Mean normalized amplitude of each group (Chinese, English) per component (Na–Pb, Pb–Nb, FFR) and pitch height stimulus (L, M). Pooling across stimulus and component, Chinese amplitudes are larger than English. By component, mid pitch height (M) is larger than low (L) for both cortical components (Na–Pb, Pb–Nb; *** $p < 0.0001$); conversely, L is larger than M for the brainstem component (FFR; *** $p < 0.0001$). The mid pitch height also evokes larger amplitude in Na–Pb than FFR; contrariwise, Pb–Nb has larger amplitude than FFR. At the cortical level, the increase in magnitude of CPR components with increasing pitch height is consistent with an increase in synchronized neural activity according to a rate-based encoding scheme. At the brainstem level, the decrease in F0 magnitude of the

FFR with increasing pitch height reflects a decrease in neural phase locking according to a temporal encoding scheme.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

F0 characteristics of pitch height stimuli (Low, Mid, High)

	Onset	Offset	Mean	Range	Minimum	Maximum
LOW	100	128	108	31	97	128
MID	190	218	198	31	187	218
HIGH	280	308	288	31	277	308

Note. F0 values are expressed in Hz, Range = F0 between minimum and maximum