Published in final edited form as:

Hear Res. 2019 June ; 377: 61–71. doi:10.1016/j.heares.2019.03.009.

Tone language experience-dependent advantage in pitch representation in brainstem and auditory cortex is maintained under reverberation

Ananthanarayan Krishnan^a, Chandan H. Suresh^{a,1}, 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

Long-term language and music experience enhances neural representation of temporal attributes of pitch in the brainstem and auditory cortex in favorable listening conditions. Herein we examine whether brainstem and cortical pitch mechanisms—shaped by long-term language experience—maintain this advantage in the presence of reverberation-induced degradation in pitch representation. Brainstem frequency following responses (FFR) and cortical pitch responses (CPR) were recorded concurrently from Chinese and English-speaking natives, using a Mandarin word exhibiting a high rising pitch ($/y\acute{r}^2/$). Stimuli were presented diotically in quiet (Dry), and in the presence of Slight, Mild, and Moderate reverberation conditions. Regardless of language group, the amplitude of both brainstem FFR (F0) and cortical CPR (Na-Pb) responses decreased with increases in reverberation. Response amplitude for Chinese, however, was larger than English in all reverberant conditions. The Chinese group also exhibited a robust rightward asymmetry at temporal electrode sites ($T8 > T7$) across stimulus conditions. Regardless of language group, direct comparison of brainstem and cortical responses revealed similar magnitude of change in response amplitude with increasing reverberation. These findings suggest that experience-dependent brainstem and cortical pitch mechanisms provide an enhanced and stable neural representation of pitch-relevant information that is maintained even in the presence of reverberation. Relatively greater degradative effects of reverberation on brainstem (FFR) compared to cortical (Na-Pb) responses suggest relatively stronger top-down influences on CPRs.

Keywords

pitch encoding; cortical pitch response (CPR); frequency following response (FFR); reverberation; spectrotemporal degradation; top-down mechanisms

Corresponding author: Jackson T. Gandour, PhD, Purdue University, Department of Speech Language Hearing Sciences, Lyles-Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907 USA, gandour@purdue.edu.

¹Present address: California State University, LA

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

1. INTRODUCTION

Human communication frequently occurs in the presence of concurrent sound sources in noisy and/or reverberant environments. These adverse listening environments pose a challenge to the auditory system's ability to extract important acoustic features of speech sounds (Sabine, 1922; Wang & Brown, 2006). In reverberant surroundings, original sound waves are reflected from walls, floors, and ceilings. Reflected sound waves arrive at the listener's ear delayed relative to the original. The temporal overlap between them results in a noisier signal compromised by multiple delays, attenuation, and spectrotemporal distortions.

Psychophysical studies have shown that reverberation-induced spectrotemporal distortion can have deleterious effects on an individual's ability to identify and discriminate acoustic features of consonants (Gelfand & Silman, 1979); vowels (Drgas & Blaszkak, 2009; Nabelek & Letowski, 1988); time-varying formant cues (Nabelek & Dagenais, 1986b) and selective attention to one talker (Culling et al., 2003; Darwin & Hukin, 2000). This degradation in identification and discrimination tasks reflects both forward masking, where preceding segments mask subsequent segments, and self-masking, where temporal smearing occurs within each phoneme (Nabelek et al., 1989; Wang & Brown, 2006). Physiological data from single neurons in the ventral cochlear nucleus (Sayles & Winter, 2008; Sayles et al., 2015), and electrophysiological data from the scalp-recorded, frequency following response (FFR) at the level of the brainstem (Bidelman, 2017; Bidelman & Krishnan, 2010) show that neural representation of pitch-relevant information based on neural phase-locking is severely degraded in the presence of reverberation. Physiologic data show that increases in reverberation smear both amplitude envelope and temporal fine structure over the duration of the stimulus (Sayles & Winter, 2008). As a consequence, pitch-related information is reduced in the temporal pattern of neural activity.

Voice pitch, a robust perceptual cue, plays an important role in music, language, speech and in the segregation of concurrent sounds (Oxenham, 2012; Plack et al., 2005). As such, measures of neural activity in the brainstem and auditory cortex permit evaluation of experience-dependent representations of pitch-relevant information at both levels of a well-coordinated, hierarchical pitch processing network.

The FFR and 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. FFR reflects sustained phase-locked neural activity in a population of neural elements primarily within the rostral brainstem (Chandrasekaran & Kraus, 2010; Krishnan, 2007; Suresh et al., 2017). Pitch-relevant neural activity preserved in the FFR is strongly correlated with behavioral measures of pitch (Bidelman & Krishnan, 2010; Krishnan & Plack, 2009; Krishnan et al., 2010b; Parbery-Clark et al., 2009). Thus, we infer that acoustic features of stimuli associated with pitch are already emerging in neural representations at the level of the brainstem. 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 POR (Gutschalk et al., 2002; Krumbholz et al., 2003)—corroborated by human depth electrode recordings (Griffiths et al., 2010;

Schonwiesner & Zatorre, 2008) and source analysis of the EEG-derived Na (Bidelman & 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). In addition, the CPR is characterized by multiple transient components (Na: 125–150 ms, Pb: 200–220, Nb: 270–285 ms). They index different temporal pitch attributes that are subject to language experience-dependent enhancement (for reviews, Gandour & Krishnan, 2016; Krishnan & Gandour, 2014).

There is compelling empirical evidence supporting the view that both language and music experience enhances neural representation of behaviorally-relevant temporal attributes of pitch in both the brainstem and auditory cortex (Gandour & Krishnan, 2014; Gandour & Krishnan, 2016; Intartaglia et al., 2016; Krishnan et al., 2012a; Skoe et al., 2015; Song et al., 2008; Wong et al., 2007; Zatorre & Gandour, 2008; Zatorre & Baum, 2012). Neuroanatomic (Saldana et al., 1996) and physiologic evidence (Suga et al., 2003) support enhanced neural representation in subcortical structures mediated by corticofugal pathways suggesting a well-coordinated interplay between brainstem and cortical processes mediating pitch processing. There is indirect evidence to suggest that experience-dependent pitch representations at cortical and subcortical levels are relatively more resistant to stimulus degradation; for example, rapid pitch change (Krishnan et al., 2015a), reduced pitch salience (Krishnan et al., 2016), and background noise (Suresh et al., 2017). Similarly, brainstem neural representation of voice fundamental frequency (F0) and the first formant frequency (F1) in musicians show greater resilience to degradative effects of noise (Parbery-Clark et al., 2009) and reverberation (Bidelman & Krishnan, 2010). In both studies, the degree of degradation in brainstem neural representation correlates well with perceptual ability under adverse listening conditions. However, it is not known if experience-dependent pitch mechanisms in the brainstem and auditory cortex are sufficiently robust and stable even in reverberant conditions that degrade the spectrotemporal cues essential for pitch representation. Also, direct comparison of concurrently-recorded brainstem and cortical pitch-relevant neural activity will allow us to determine whether pitch mechanisms at the brainstem and cortical levels are differentially susceptible to reverberation-induced degradation.

In this study, FFRs and CPRs were recorded concurrently from Mandarin- and English-speaking natives as they listened passively to a high rising lexical tone (/yí²/) superimposed on a Mandarin syllable. Stimuli were presented diotically in Quiet, and in the presence of Slight, Mild, and Moderate reverberation. We hypothesize that both groups will show degradation in pitch-relevant information under reverberant conditions at both brainstem and cortical levels. However, more robust and stable language experience-shaped neural mechanisms in the brainstem and auditory cortex are hypothesized to maintain the experience-dependent advantage in pitch representation even under reverberation. A direct comparison of cortical and brainstem responses is expected to reveal differences in the patterns of reverberation-induced degradation, pointing to differential susceptibility of these responses to degradative effects of reverberation.

2. Materials and methods

2.1. Participants

EEG data were recorded from 14 native speakers of Mandarin Chinese (C: 7 male, 7 female) and English (E: 6 male, 8 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 (C: $M = 23.46$ years, $SD = \pm 2.7$; E: $M = 21.62$ years, $SD = \pm 1.9$), and years of formal education (C: $M = 17.1$ years, $SD = \pm 2.6$; E: $M = 16.1$ years, $SD = \pm 1.63$). They were strongly right handed (C: $M = 89.0\%$, $SD = \pm 18.1\%$; E: $M = 98.0\%$, $SD = \pm 6.93\%$) as measured by the laterality index of the Edinburgh Handedness Inventory (Oldfield, 1971; Somers et al., 2015). All Chinese participants were born and raised in mainland China. Only three Chinese participants had received formal instruction in English before the age of ten ($M = 12$ years, $SD = \pm 2.80$ 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, 4.6 and 5.1, respectively (Li et al., 2006). Their daily usage of Mandarin and English, in order, was reported to be 72% and 28%. As determined by a music history questionnaire (Wong & Perrachione, 2007), all Chinese ($M = 0.79$ years, $SD = \pm 0.77$) and English ($M = 0.79$ years, $SD = \pm 0.86$) participants had no more than two years of musical training on any combination of instruments. No participant had any training within the past five years. Each participant was paid and gave informed consent in conformity with the 2013 World Medical Association Declaration of Helsinki and in compliance with a protocol approved by the Institutional Review Board of Purdue University.

2.2. Stimuli

Four stimulus conditions were constructed to evaluate the degradative effects of reverberation on the representation of pitch-relevant information at the level of the brainstem auditory cortex. The original stimulus condition (Dry) was an unaltered version (i.e., no reverb) of the Mandarin speech syllable (/yí²/). It exhibited an F0 trajectory exemplary of the high rising Tone 2 (Xu, 1997; Fig. 8C, top panel, white curvature). Mean F0 was 111 Hz; onset and offset frequencies were 102 Hz and 130 Hz, respectively. Both fundamental frequency and formants (onset: F1 = 306 Hz, F2 = 2448 Hz, and F3 = 3672 Hz) were time-varying. The addition of reverberation was accomplished by time-domain convolution of the Dry stimulus with room impulse responses recorded in a corridor at a distance from the sound source of 0.63 m (Slight: reverberation time (RT60) ≈ 0.7 s); 1.25 m (Mild: RT60 ≈ 0.8 s); and 5 m (Moderate: RT60 ≈ 0.9 s) (Sayles & Winter, 2008; Watkins, 2005). We removed the 'tails' and silent intervals introduced by convolution. All stimulus waveforms were normalized to a duration of 250 ms and equal RMS amplitude. This normalization rules out overall amplitude as a confounding factor (Sayles & Winter, 2008). Figure 1 shows clear periodicity (time waveform; left panel) and pitch contour (spectrogram; right panel) in the Dry stimulus. With increasing levels of reverberation (Slight, Mild, Moderate), we observe considerable degradation of waveform periodicity and pitch contour.

We adopted Krumbholz et al.'s (2003) novel stimulus paradigm to extract pitch-specific CPR components. In order to disentangle the pitch-specific response from the obligatory onset responses (P1-N1-P2), they constructed a continuous sound using an initial segment of noise with no pitch followed by a pitch-eliciting segment matched in intensity and overall spectral profile. The initial segment of noise evokes only the obligatory components. Interestingly, the pitch-specific response is evoked for the noise-to-pitch transition but not for the pitch-to-noise stimulus transition (Krishnan et al., 2012b; Krumbholz et al., 2003). Specifically, the experimental paradigm consisted of two segments: a 250 ms pitch segment preceded by an 800 ms noise segment. The noise segment is spectrally matched to the pitch-eliciting stimulus, giving it the percept of whispered/i/. This spectral matching was achieved by filtering the noise using linear prediction filter coefficients of the speech segment.

The overall RMS level of each segment was equated such that a continuous stimulus was created with no discernible difference in amplitude among the two segments. The noise segments were crossfaded with the pitch segment using 7 ms cos^2 ramps to ensure no amplitude differences at the noise-to-pitch transition. This resulted in an overall stimulus duration of 1043ms. The overall level of each combined stimulus was 80 dB SPL. All stimulus conditions were presented through magnetically-shielded tubal insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA) at a repetition rate of 0.56/s using a fixed onset polarity (rarefaction) for the pitch-eliciting stimulus. Stimulus presentation order was randomized both within and across participants. All stimuli were generated and played out using an auditory evoked potential system (SmartEP, Intelligent Hearing Systems; Miami, FL, USA).

2.2.1. 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 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. 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. 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. 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 canthus of the left eye and another electrode placed above the left eye was used to monitor artifacts introduced by ocular activity. Inter-electrode impedances were maintained below 10 k Ω . For each stimulus condition, EEGs were acquired in two sessions of 750 sweeps each. The experimental protocol took about 2 hours to complete.

2.2.2. Extraction of CPR and FFR—CPR and FFR responses were extracted off-line from the EEG files using different re-referenced electrode montages. We analyzed CPR data on 14 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. Criteria for exclusion included poor repeatability of the FFR waveforms, artifact rejection greater than 10%, and absence of clearly discernible spectral peaks at F0. They were replaced by FFR data from three new participants for each group. This new data set ($n = 14$) was used to characterize 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. EEG data were then digitally band-pass filtered (2–25 Hz, Butterworth zero phase shift filter with 24 dB/octave rejection rate) to enhance the transient components and minimize the sustained component. Sweeps containing electrical activity exceeding $\pm 75 \mu\text{V}$, including ocular artifacts, were rejected automatically. Subsequently, averaging was performed on all unipolar electrode locations. Using the common reference, we were able to compare CPR components at the right (T8) and left (T7) temporal electrode sites for hemispheric asymmetry effects. Given the poor spatial resolution of EEG even using multiple electrodes, the focus here was not to localize the source of the CPRs with just two electrodes. Instead, our aim was 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., 2015b; Krishnan et al., 2014b; Krishnan et al., 2015c). The re-referenced electrode site, Fz-linked (T7/T8), was used to characterize the transient pitch response components. It was chosen because both MEG- and EEG-derived pitch responses are prominent at frontocentral sites (e.g., Bidelman & Grall, 2014; Krishnan et al., 2015b; Krumbholz et al., 2003). It also allowed us to compare our CPR data with Fz-derived POR data (Bidelman & Grall, 2014; Gutschalk et al., 2002; Gutschalk et al., 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 and (ii) the inverted components recorded at the temporal electrode sites (T7 and T8). For both averaging procedures, the analysis epoch was 1500 ms, including the 100 ms pre-stimulus baseline.

To extract the much smaller amplitude of the brainstem FFR, EEG data were digitally band-pass filtered (75–1500 Hz, Butterworth zero phase filters with 24 dB/Octave rejection rate). Sweeps containing electrical activity exceeding $\pm 40 \mu\text{V}$ 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; 743 represents the onset of the pitch segment). Each FFR waveform represents the grand average of the FFRs derived from the three electrode montages in response to a total of 1500 sweeps obtained in two recordings of 750 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 signal SNR and detectability of the FFR, particularly the degraded responses under reverberation.

2.2.3. CPR magnitude—The evoked response to the entire two-segment stimulus paradigm (noise-pitch) is characterized by obligatory components (P1/N1) corresponding to the onset of energy in the precursor noise segment. They are followed by three transient CPR components (Na, 135–170 ms; Pb, 210–240 ms; Nb, 270–310 ms) occurring after the onset of the pitch-eliciting segment. We evaluated only the magnitude of the CPR components. Peak-to-peak amplitude of Na–Pb and Pb–Nb were measured manually to characterize the effects of reverberation on the pitch-relevant activity reflected in the CPR components. Manual peak picking was necessary due to decreased neural reliability of automatic peak picking associated with small, degraded response waveform morphology with added reverberation. To improve accuracy and consistency of manual peak picking for latency and amplitude measurements from individual averaged responses, the grand-averaged response was used as a reference 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 (95%).

For each condition, peak-to-peak amplitude of Na–Pb was measured at the temporal electrode sites (T7/T8) to evaluate hemispheric asymmetry using the laterality index (LI). $LI = (Na-Pb_{T8} - Na-Pb_{T7} / Na-Pb_{T8} + Na-Pb_{T7})$, where Na–Pb_{T8} and Na–Pb_{T7} refers to the amplitude of the Na–Pb response at T8 and T7 electrode sites, respectively. Positive LI values reflect rightward asymmetry; negative LI values reflect leftward asymmetry. To enhance visualization of hemispheric asymmetry effects along a spectrotemporal dimension, a joint time frequency analysis using a continuous wavelet transform was performed on the grand average waveforms derived from the temporal electrodes.

2.2.4. FFR neural periodicity strength—Neural periodicity strength was quantified by measuring the magnitude of the F0 component for two temporal windows of the response (one over the initial slow changing pitch (10–100 ms), and the other over the rapidly changing pitch segment (100–225ms) from each response waveform. The spectrum of each response segment was computed by obtaining the Fast Fourier Transform (FFT) of time-windowed version of its temporal waveform (Gaussian window, 1Hz resolution). For each subject, the magnitude of F0 was measured as the peak in the FFT relative to the noise floor. This floor was estimated by averaging the amplitude values over a 50 Hz band immediately below and above the F0 peak in the spectrum. All FFR data analyses were performed using custom routines coded in Matlab16 (The Math Works, Inc., Natick, MA, USA). To enhance the visualization of spectrotemporal changes in the FFR, a joint time frequency analysis using a continuous wavelet transform was performed on the grand average waveforms.

2.2.5. Comparison of CPR and FFR—Comparison of the CPR and FFR data was limited to the original 11 subjects (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. Therefore, to

allow meaningful comparison of the CPR and FFR responses, we normalized the change in magnitude in reference to the Dry condition [(Dry-Slight)/Dry; (Dry-Mild)/Dry; and (Dry-Moderate)/Dry].

2.3. Statistical analysis

To evaluate the effects of reverberation, a two-way (group x stimulus) mixed model ANOVAs (SAS®; SAS Institute, Inc., Cary, NC, USA) was performed on the peak-to-peak amplitude of Na–Pb derived from the Fz electrode site. Language group (Chinese, English) served as the between-subjects factor; subjects nested within group as the random factor. Stimulus condition (Dry, Slight, Mild, Moderate) served as the within-subjects factor. To assess asymmetry in CPR response amplitudes at temporal electrode sites (T7 [left], T8 [right]), a two-way (group x stimulus) ANOVA was performed on the laterality index (T8–T7/T8+T7) for each stimulus condition. To evaluate the FFR, a two-way (group x stimulus) ANOVA was performed on the magnitude of peak F0 within two 40 ms regions (10–100; 100–225) that contrast in rate of acceleration (slow, fast). In the rapid-changing pitch region, data were transformed to a $\log_{10}$ scale in order to satisfy normality and homogeneity of variance assumptions. To compare changes in response magnitude between reverberant conditions at both cortical (Na–Pb) and brainstem levels of representation (FFR slow, FFR fast), a two-way (component x stimulus) was performed on normalized data including Dry-Slight, Dry-Mild and Dry-Moderate conditions. To justify pooling across groups for purposes of comparing cortical versus brainstem effects, a preliminary ANOVA (group x stimulus) revealed no language group effects (Na–Pb: $F_{1,18} = 0.36$, $p = 0.5562$; FFR slow: $F_{1,18} = 0.01$, $p = 0.9124$; FFR fast: $F_{1,18} = 0.03$, $p = 0.8617$).

Normality of distribution (Kolmogorov-Smirnov) and homogeneity of residual variance (Levene) assumptions were confirmed prior to statistical inference. All post hoc multiple comparisons were corrected with a Bonferroni significance level set at $\alpha = 0.05$.

Partial eta squared (η_p^2) values were reported to indicate effect sizes.

3. Results

3.1. Cortical pitch response component

3.1.1. Fz: Response morphology of CPR components—Fig. 2 shows grand-averaged cortical pitch response waveforms elicited by the Dry and reverberant conditions for the Chinese (red) and the English (blue) group. CPR components (gray background) are clearly identifiable in both groups. Regardless of group, amplitude of pitch-relevant components (Na, Pb, Nb) appear to decrease in amplitude with increased reverberation. Pitch components, however, are clearly more robust for the Chinese group irrespective of reverberant conditions. Their overall better morphology may indicate that the experience-dependent advantage in pitch representation is maintained even in the presence of the degradative effects of reverberation.

3.1.2. Fz: Response magnitude of CPR (Na–Pb)—Mean peak-to-peak amplitude of Na–Pb elicited by the Dry and reverberant stimuli are displayed in Fig. 3. ANOVA revealed

main effects of group ($F_{1,24} = 37.97, p < 0.0001, \eta_p^2 = 0.613$) and stimulus ($F_{3,72} = 36.17, p < 0.0001, \eta_p^2 = 0.601$). Na–Pb amplitude of the Chinese group was larger than that for the English group ($C > E$) across reverberant conditions. Response amplitude showed a progressive degradation effect with increasing strength of reverberation (Dry > Slight > Mild > Moderate). Dry was larger than Slight, Mild, and Moderate; Slight was larger than Mild and Moderate; and Mild was larger than Moderate. The relatively greater Na–Pb amplitude for the Chinese in reverberant conditions suggest language experience enhances resilience of pitch-relevant neural activity in the auditory cortex to reverberation-induced degradation.

3.1.3. T7/T8: laterality index of CPR component (Na–Pb)—Fig. 4 shows the laterality index of Na–Pb at the temporal electrode sites (T7, T8) for each language group and reverberant condition. ANOVA revealed a main effect of group only ($C > E$: $F_{1,24} = 45.64, p < 0.0001, \eta_p^2 = 0.655$). Neither the main effect of stimulus ($F_{3,72} = 2.46, p = 0.0694, \eta_p^2 = 0.093$) nor the interaction of group and stimulus reached significance ($F_{3,72} = 0.83, p = 0.4808, \eta_p^2 = 0.033$). Pooled across reverberant conditions, a strong rightward asymmetry (T8 > T7) was found in the Chinese group. No strong asymmetries were evident in the English group. The Chinese group's rightward asymmetry indicates an experience-dependent preferential recruitment of the right auditory cortex for pitch processing.

3.2. Brainstem response component

3.2.1. Spectrotemporal response characteristics of FFR—Fig. 5 displays grand averaged FFR waveforms (left: Chinese listeners, red; English, blue) and spectrograms (right) for the Dry and reverberant conditions. FFR waveforms show that their magnitude decreases with increasing reverberation across groups. Yet the Chinese group's waveforms appear to be more robust than those of the English. Chinese listeners' spectrograms—in comparison to those of English listeners—show a relatively more robust band of phase-locking following the rising F0 contour for Dry and reverberant conditions (Slight, Mild, Moderate). The magnitude of the F0 band of phase-locking representing the later, fast-acceleration pitch segment (red bar) also appears to decrease with increasing reverberation—especially Mild and Moderate—more dramatically for the English compared to the Chinese group. In those reverberant conditions, the spectrogram obtained for the English group shows a flat, smeared band of activity over the fast-acceleration pitch segment.

3.2.2. Neural periodicity strength of FFR as a function of reverberant condition—Fig. 6 shows mean F0 magnitude of FFR encoding at each step along the reverberation continuum for slow- (top panel) and fast-acceleration (bottom panel) segments of the pitch contour. ANOVA performed on F0 magnitude data for the slow-acceleration segment revealed significant main effects of group ($F_{1,20} = 167.15, p < 0.0001, \eta_p^2 = 0.893$) and stimulus ($F_{3,60} = 18.71, p < 0.0001, \eta_p^2 = 0.825$). The group $\times$ stimulus interaction was not significant. For the Chinese group, simple main effects of stimulus showed that Dry ($p = 0.0003$) and Slight ($p < 0.0001$) were larger in F0 magnitude than Mild as well as Moderate (Dry, $p < 0.0001$; Slight, $p < 0.0001$). For the English group, only Dry was larger in F0

magnitude than Mild ($p = 0.0075$) and Moderate ($p = 0.0008$). Per stimulus condition, F0 magnitude of the Chinese group was larger than that of the English group across the board (Dry, $p < 0.0001$; Slight, $p < 0.0001$, Mild, $p < 0.0001$; Moderate, $p = 0.0002$). ANOVA performed on normalized F0 magnitude ($\log_{10}$) data for the fast-acceleration segment similarly revealed significant main effects of group ($F_{1,20} = 66.71$, $p < 0.0001$, $\eta_p^2 = 0.769$) and stimulus ($F_{3,60} = 39.73$, $p < 0.0001$, $\eta_p^2 = 0.665$). The group $\times$ stimulus interaction was not significant. For the Chinese group, simple main effects of stimulus showed that Dry and Slight were larger in F0 magnitude than either Mild (Dry, $p < 0.0014$; Slight, $p < 0.0115$) or Moderate (Dry, $p < 0.0001$; Slight, $p < 0.0001$). In addition, F0 magnitude of Mild was larger than Moderate ($p = 0.0020$). The English group showed that Dry was larger in F0 magnitude than Mild ($p = 0.0014$) or Moderate ($p < 0.0001$); Slight and Mild were larger than Moderate, ($p = 0.0004$) and ($p = 0.0329$), respectively. In contrast to the Chinese group, the difference in F0 magnitude between Slight and Mild failed to reach significance. Per stimulus condition, F0 magnitude of the Chinese group was larger than that of the English group across the board (Dry, $p < 0.0001$; Slight, $p < 0.0001$, Mile, $p < 0.0001$; Moderate, $p = 0.0002$). F0 magnitude of the Chinese group was larger than English. Regardless of group, F0 magnitude for the Dry condition was larger than Mild or Moderate. The F0 magnitude of the Slight condition was larger than Mild for the Chinese group only. The language-dependent advantage appears to be maintained across reverberant conditions irrespective of changes in rate of acceleration.

3.3. Comparison of the effects of reverberation on brainstem and cortical responses

Fig. 7 displays normalized amplitude changes—relative to the Dry condition—for the cortical Na–Pb component of the CPR (top panel), FFR F0 magnitude changes to the slow-acceleration pitch segment (middle panel), and FFR F0 magnitude changes to the fast-acceleration pitch segment (bottom panel). Pooled across language groups, results of a two-way ANOVA (component $\times$ stimulus) on normalized change in response amplitude yielded a significant main effect of stimulus only ($F_{2,114} = 57.47$, $p < 0.0001$, $\eta_p^2 = 0.502$). The change in normalized magnitude from the Dry-Slight condition was smaller than Dry-Mild or Dry-Moderate. In addition, the amplitude change for the Dry-Mild condition was smaller than Dry-Moderate. Irrespective of pitch component—cortical or brainstem—response amplitude change gets larger with increasing reverberation. Two-way ANOVAs (group $\times$ stimulus) were then carried out on *normalized* change in response magnitude per component. ANOVAs yielded a significant main effect of stimulus only (CPR Na–Pb: $F_{2,36} = 15.71$, $p < 0.0001$, $\eta_p^2 = 0.169$; FFR slow-acceleration segment: $F_{2,36} = 27.32$, $p < 0.0001$, $\eta_p^2 = 0.603$; FFR fast-acceleration segment: $F_{2,36} = 18.16$, $p < 0.0001$, $\eta_p^2 = 0.502$). For the cortical component (CPR Na–Pb), post hoc comparisons showed that the amount of change in magnitude for the Dry-Slight condition was smaller than Dry-Mild ($p < 0.0102$) and Dry-Moderate ($p < 0.0001$). For both FFR slow- and fast-acceleration components, post hoc comparisons showed not only that the amount of change in magnitude for the Dry-Slight condition was smaller than Dry-Mild (slow-acceleration, $p = 0.0119$; fast-acceleration, $p = 0.0050$) and Dry-Moderate (slow-acceleration, $p < 0.0001$; fast-acceleration, $p = 0.0045$) conditions, but also that Dry-Mild was also smaller than Dry-Moderate (slow-acceleration, p

= 0.0004; fast-acceleration, $p = 0.0045$). These data suggest that sensitivity to reverberant effects on pitch processing varies depending on the level of brain structure. Herein, the brainstem response (FFR) is more sensitive than the cortical (Na–Pb). These aggregate data suggest that the sustained fine-grained representation of spectrotemporal attributes of pitch reflected in the brainstem is more susceptible to degradation when compared to an integrated estimate of pitch and its strength reflected in the auditory cortex.

4. Discussion

This study demonstrates that neural representation of pitch-relevant information in the brainstem (FFR) and auditory cortex (CPR) shows progressive deterioration with increasing reverberation for both Chinese and English listeners as a consequence of increasing deterioration in spectrotemporal distortion of acoustic cues important for pitch. Nevertheless, enhanced and more stable pitch-relevant responses in the Chinese—by virtue of long-term native language experience—help to sustain the advantage even in reverberant conditions at both cortical and brainstem levels of processing. At the cortical level, we also observed a rightward asymmetry across stimulus conditions for the Chinese group only (T8 > T7), suggesting an experience-dependent recruitment of more optimal resources for pitch processing in the right hemisphere. A direct comparison of cortical and brainstem responses reveals that the magnitude of change in the pitch responses at the brainstem and cortical levels are comparable across groups. This finding suggests that pitch mechanisms in both Chinese and English listeners are equally susceptible to degradative effects of reverberation. The relatively greater effects of reverberation on brainstem responses vis à vis cortical responses imply differential susceptibility of pitch processing to reverberation at these two levels of pitch representation.

4.1. Neural bases of reverberation-induced degradation in brainstem and cortical pitch representation

For both groups, increasing reverberation progressively reduced the amplitude of both brainstem and cortical pitch responses, a likely consequence of increasingly greater degradation of spectrotemporal pitch cues in the stimulus. With increasing reverberation, the reduced amplitude of CPR components reflect neural desynchronization of pitch-related neural activity due to disrupted coherence (Koerner & Zhang, 2015). Such desynchronization likely manifests as spectrotemporal smearing of acoustic cues relevant to pitch. This results in a widening of the temporal window to integrate pitch information across frequency channels. This wider temporal window, in turn, results in the degraded neural representation of pitch, and pitch strength (Na–Pb) as well as dynamic temporal attributes of pitch (Pb–Nb). The more severe degradation of the later Pb–Nb component precluded measurement and analysis. The more severe degradation of the later Pb–Nb component did not allow us to reliably measure its amplitude, and therefore was excluded from analysis. We have previously demonstrated that CPR components appear to be differentially sensitive to specific features of dynamic pitch (Krishnan et al., 2014a; Krishnan et al., 2015b; Krishnan et al., 2017a; Krishnan et al., 2017b; Krishnan et al., 2014b). The later Pb–Nb time window optimally represents pitch acceleration (Krishnan et al., 2015a). Reverberation has been shown to produce a more pronounced temporal smearing

of the time-varying features of speech sounds (Nabelek & Dagenais, 1986a). It is plausible that Pb-Nb, a temporal processing window that indexes extraction of the more rapidly accelerating portion of the pitch contour, would be more susceptible to reverberation-induced degradation compared to a temporal window indexing the earlier essentially flat portion of the pitch contour (Na-Pb).

In this study, the brainstem FFRs also show progressive degradation of the neural representation of pitch with increasing reverberation. Our findings are consistent with animal studies that show neural encoding of pitch based on timing information is severely degraded in the presence of reverberation in the caudal brainstem (Sayles & Winter, 2008; Sayles et al., 2015, ventral cochlear nucleus). Reverberation degrades the sustained fine-grained, phase-locked temporal discharge pattern of brainstem neural responses following the pitch contour. This degradation is likely due to the loss of robust periodicity and smearing of dynamic changes over time in the fine structure of the stimulus (Houtgast & Steeneken, 1973; Houtgast & Steeneken, 1985; Plomp & Steeneken, 1973). In the presence of reverberation, we infer that degraded neural representation of pitch at the cortical level reflects, in part, degraded brainstem inputs driving the neural generators of the CPR.

For cortical and brainstem responses, no differences between language groups were revealed in comparisons of change in amplitude with increasing reverberation. These findings suggest that pitch encoding mechanisms and their susceptibility to adverse listening conditions are fundamentally similar irrespective of language typology. Nonetheless, we observe more robust pitch-relevant neural activity for Chinese listeners at both brainstem and cortical levels even in reverberant conditions. This suggests that long-term experience shapes pitch mechanisms in order to preserve a gain in signal-to-noise ratio in the presence of reverberation. Such findings add to the extant literature on experience-dependent, compensatory adjustment to deleterious effects of adverse listening conditions on both subcortical and cortical representation of complex sounds (Bidelman, 2018; Bidelman et al., 2018; Krishnan et al., 2010a; Parbery-Clark et al., 2009; Suresh et al., 2017).

4.2. Experience-dependent enhancement in pitch representation is maintained under reverberant conditions

At the frontotemporal electrode site (Fz), the Chinese group exhibited greater Na-Pb amplitude than the English group in Dry and all reverberant conditions. This finding is consistent with our previous reports that have consistently shown experience-dependent enhancement of Na-Pb with manipulation of several temporal attributes of pitch (Bidelman & Krishnan, 2010; Krishnan et al., 2017b; and references therein). The fact that the Chinese advantage is maintained under adverse conditions suggests that cortical pitch mechanisms shaped by experience are not only more stable but also better able to extract pitch-relevant information from degraded, linguistically-relevant acoustic features.

Maintaining experience-dependent enhancement is essentially an increase in signal-to-noise ratio of pitch-relevant neural activity. It is crucial in preserving behaviorally-relevant acoustic features that are degraded in adverse listening conditions (Bidelman et al., 2018; Pantev et al., 2001; Rammsayer & Altenmüller, 2006). Chinese listeners' enhancement of neural representation in reverberant conditions as well as Dry is likely mediated by selective

recruitment of neural elements with greater temporal synchronization. In our view, the Chinese group's enhancement represents modulation of early sensory pitch mechanisms in the brainstem and auditory cortex that interact with extrasensory processes at a higher hierarchical level in the auditory cortex (see also Krishnan et al., 2016, p. 111). Such adaptive pitch mechanisms are also essential to sharpen neural response properties to optimally capture temporal attributes of pitch in reduced SNR conditions (Suresh et al., 2017).

Our inability to reliably detect and analyze the later Pb–Nb component of the CPR suggests greater susceptibility of this component to reverberation. The later Pb–Nb time window optimally represents pitch acceleration (Krishnan et al., 2015a). In this study, Pb–Nb indexes extraction of the later fast-acceleration segment of the pitch contour. Reverberation is known to produce a pronounced, temporal smearing of time-varying features of speech sounds (Nabelek & Dagenais, 1986a). Pb–Nb may be more susceptible to reverberation-induced degradation compared to a temporal window indexing the earlier slow-acceleration segment of the pitch contour (Na–Pb). Consistent with this view, the later integration window (Pb–Nb) has also been shown to be more susceptible to noise degradation (Suresh et al., 2017). In fact, no experience-dependent effect was observed in the presence of background noise at 0 dB SNR. Thus, relative weighting of the dynamic, temporal attributes of pitch varies depending on both language experience and sensitivity/stability of pitch-relevant neural activity within a given temporal integration window. These findings suggest that robust, experience-shaped pitch mechanisms are ineffective if spectrotemporal cues are severely degraded by either noise or reverberation.

A qualitative comparison of changes in the CPR due to reverberation (herein) and background noise (Suresh et al., 2017, Fig. 3) reveal that both types of adverse listening conditions result in degradation of CPR neural activity. In reverberation, language-dependent enhancement is maintained across conditions; in background noise, it is maintained only for the most favorable SNR (+5 dB) condition. Taken together, these results are consistent with Bidelman et al. (2018). They demonstrated that speech-evoked brainstem and early cortical responses were degraded more by noise than reverberation. Previous physiological studies (Al Osman et al., 2017; Bidelman & Krishnan, 2010; Sayles & Winter, 2008) have also shown better preservation of neural representation in reverberation compared to noise (Bidelman, 2017, review). It remains unclear if these differences simply reflect fundamental differences in the extent to which spectrotemporal pitch-relevant acoustic features are disrupted and/or the relative effectiveness of masking mechanisms under both noise and reverberation. We cannot rule out interplay between bottom up and top down mechanisms (Suga et al., 2000; Suga et al., 2003) to aid signal selection in adverse listening conditions. Such interplay, however, does not explain why signal selection would be selectively aided in reverberation but not noise. Background noise more closely matches the spectrum of the speech signal. Consequently, it may be a more effective masker of acoustic features over the entire duration of the stimulus. In reverberation, acoustic cues relevant to pitch are likely to remain available for at least some portions of the signal.

4.3. Experience-dependent preferential recruitment of the right hemisphere is maintained under reverberant conditions

At the Fz frontotemporal electrode site and Fz-linked T7/T8 temporal electrode configuration, Na–Pb amplitude of the Chinese group is larger than that of the English, pooled across reverberant conditions. With respect to hemispheric preferences, it is only over the right temporal site (T8) that amplitude of the Chinese group is larger than that of the English. These data are consistent with recent reports of rightward asymmetry for other temporal attributes of pitch: rate of pitch acceleration (Krishnan et al., 2015a); pitch direction and location of peak acceleration (Krishnan et al., 2015b); location of turning point (Krishnan et al., 2017c); changes in pitch height (Krishnan et al., 2017b). Overall, these findings converge with an extant literature that highlights the role of the right hemisphere in processing both linguistic and musical pitch (for reviews, Friederici & Alter, 2004; Gandour & Krishnan, 2016; Kwok et al., 2017; Zatorre & Gandour, 2008; Zatorre & Baum, 2012; Zatorre et al., 2002).

4.4. Pitch processing at the brainstem and cortical levels show differential susceptibility to reverberation

Comparison of reverberation-induced changes in response magnitude with increasing reverberation reveals a relatively greater reduction for the brainstem FFR compared to the cortical Na–Pb. One possible explanation is that sustained phase-locking is essential to capture the fine-grained details of time-varying acoustic properties relevant to pitch. Thus, FFR is more susceptible to degradative effects of reverberation compared to the temporally integrated, less fine-grained, pitch-relevant activity reflected in the CPR. It is also possible that the effectiveness of top-down, facilitatory influence may differ on cortical and brainstem pitch processes (Bidelman et al., 2018). Their speech data show larger reverberation-related changes in the brainstem FFR relative to the cortical N1 response. They infer that stronger extrasensory, top-down, facilitatory influence on sensory level responses generated in the primary auditory cortex (Galbraith & Kane, 1993; Hillyard & Picton, 1979; Okamoto et al., 2011) are stronger than cortico-collicular influence on brainstem FFRs. However, the fact that long-term language experience enhances and maintains the advantage in reverberation for the Chinese listeners at both levels does not rule out local brainstem and/or cortico-collicular facilitatory influence in adverse listening conditions. Taken together, these results suggest the operation of a well-coordinated hierarchical network for pitch processing.

5. Conclusions

Effects of reverberation on neural representation of pitch-relevant information in the brainstem and auditory cortex reveal changes that are shaped not only by language-independent auditory processes—reflected by similar amplitude change with increasing reverberation across groups—but also by higher-order processes shaped by long-term language experience—reflected by maintenance of enhanced responses and rightward asymmetry in dry and reverberant conditions). We infer that long-term language experience shapes and strengthens adaptive, distributed hierarchical pitch processing along the auditory pathway. Experience-dependent effects reflect local and top-down modulation of early sensory level processing of pitch-relevant information in the auditory cortex. Such

enhancements enable more robust and stable representation of behaviorally-relevant temporal pitch attributes in both normal and adverse listening conditions. Differential effects of reverberation on brainstem and cortical pitch responses suggest a stronger facilitatory influence of top-down mechanisms on cortical pitch processing.

Acknowledgements

This work was supported by the National Institutes of Health 5R01DC008549–8 (A.K.). Thanks to Rongrong Zhang and Whitney Lyle for their assistance with statistical analysis and data collection, respectively.

References

- Al Osman R, Dajani HR, Giguere C 2017 Self-masking and overlap-masking from reverberation using the speech-evoked auditory brainstem response. *J. Acoust. Soc. Am* 142, EL555 10.1121/1.5017522 [PubMed: 29289064]
- Bidelman GM 2017 Communicating in challenging environments: Noise and reverberation In: Kraus N, Anderson T, White-Schwoch T, Fay RR, Popper AN, (Eds.), *Springer Handbook of Auditory Research: The frequency-following response: A window into human communication*, Vol. 61 Springer Nature, New York, NY pp. 193–224.
- Bidelman GM 2018 Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage* 175, 56–69. 10.1016/j.neuroimage.2018.03.060 [PubMed: 29604459]
- Bidelman GM, Krishnan A 2010 Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res* 1355, 112–125. [PubMed: 20691672]
- Bidelman GM, Grall J 2014 Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *Neuroimage* 101, 204–14. 10.1016/j.neuroimage.2014.07.005 [PubMed: 25019679]
- Bidelman GM, Davis MK, Pridgen MH 2018 Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. *Hear. Res* 367, 149–160. 10.1016/j.heares.2018.05.018 [PubMed: 29871826]
- Chandrasekaran B, Kraus N 2010 The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology* 47, 236–46. PSYP928 [pii] 10.1111/j.1469-8986.2009.00928.x [PubMed: 19824950]
- Culling JF, Hodder KI, Toh CY 2003 Effects of reverberation on perceptual segregation of competing voices. *J. Acoust. Soc. Am* 114, 2871–2876. [PubMed: 14650021]
- Darwin CJ, Hukin RW 2000 Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention. *J. Acoust. Soc. Am* 108 335–342. [PubMed: 10923896]
- Drgas S, Blaszk MA 2009 Perceptual consequences of changes in vocoded speech parameters in various reverberation conditions. *J. Speech. Lang. Hear. Res* 52, 945–955. [PubMed: 19380609]
- Friederici AD, Alter K 2004 Lateralization of auditory language functions: a dynamic dual pathway model. *Brain Lang* 89, 267–76. 10.1016/S0093-934X(03)00351-1 S0093934X03003511 [pii] [PubMed: 15068909]
- Galbraith GC, Kane JM 1993 Brainstem frequency-following responses and cortical event-related potentials during attention. *Percept. Mot. Skills* 76, 1231–41. 10.2466/pms.1993.76.3c.1231 [PubMed: 8337070]
- Gandour JT, Krishnan A 2014 Neural bases of lexical tone In: Winkler H, Padakannaya P, (Eds.), *Handbook of South and Southeast Asian psycholinguistics* Cambridge University Press, Cambridge, UK pp. 339–349.
- Gandour JT, Krishnan A 2016 Processing tone languages In: Hickok G, Small SL, (Eds.), *Neurobiology of language* Academic Press, New York pp. 1095–1107.
- Gelfand SA, Silman S 1979 Effects of small room reverberation upon the recognition of some consonant features. *J. Acoust. Soc. Am* 66, 22–29.

- Griffiths TD, Kumar S, Sedley W, Nourski KV, Kawasaki H, Oya H, Patterson RD, Brugge JF, Howard MA 2010 Direct recordings of pitch responses from human auditory cortex. *Curr. Biol* 20, 1128–32. 10.1016/j.cub.2010.04.044 [PubMed: 20605456]
- Gutschalk A, Patterson RD, 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, 207–16. 10.1006/nimg.2001.0949 [PubMed: 11771990]
- Gutschalk A, Patterson RD, Scherg M, Uppenkamp S, Rupp A 2004 Temporal dynamics of pitch in human auditory cortex. *Neuroimage* 22, 755–66. 10.1016/j.neuroimage.2004.01.025 S1053811904000680 [pii] [PubMed: 15193604]
- Hillyard SA, Picton TW 1979 Event-related brain potentials and selective information processing in man In: Desmedt JE, (Ed.), *Progress in Clinical Neurophysiology* Karger, Basel pp. 1–52.
- Houtgast T, Steeneken HJM 1973 The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acustica* 28, 66–73.
- Houtgast T, Steeneken HJM 1985 A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *J. Acoust. Soc. Am* 77, 1069–1077. 10.1121/1.392224
- Intartaglia B, White-Schwoch T, Meunier C, Roman S, Kraus N, Schon D 2016 Native language shapes automatic neural processing of speech. *Neuropsychologia* 89, 57–65. 10.1016/j.neuropsychologia.2016.05.033 [PubMed: 27263123]
- Johnsrude IS, Penhune VB, Zatorre RJ 2000 Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–63. [PubMed: 10611129]
- Koerner TK, Zhang Y 2015 Effects of background noise on inter-trial phase coherence and auditory N1-P2 responses to speech stimuli. *Hear. Res* 328, 113–9. 10.1016/j.heares.2015.08.002 [PubMed: 26276419]
- Krishnan A 2007 Human frequency following response In: Burkard RF, Don M, Eggermont JJ, (Eds.), *Auditory evoked potentials: Basic principles and clinical application* Lippincott Williams & Wilkins, Baltimore pp. 313–335.
- Krishnan A, Plack CJ 2009 Auditory brainstem correlates of basilar membrane nonlinearity in humans. *Audiol. Neurotol* 14, 88–97. 000158537 [pii] 10.1159/000158537 [PubMed: 18827479]
- Krishnan A, Gandour JT 2014 Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: electrophysiological evidence. *Acoustics Australia* 42, 166–178. [PubMed: 25838636]
- Krishnan A, Gandour JT, Bidelman GM 2010a Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Res* 1313, 124–33. 10.1016/j.brainres.2009.11.061 [PubMed: 19961835]
- Krishnan A, Bidelman GM, Gandour JT 2010b Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hear. Res* 268, 60–6. 10.1016/j.heares.2010.04.016 [PubMed: 20457239]
- Krishnan A, Gandour JT, Bidelman GM 2012a Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport* 23, 498–502. 10.1097/WNR.0b013e328353764d [PubMed: 22495037]
- Krishnan A, Gandour JT, Suresh CH 2014a Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. *Brain Lang* 138, 51–60. 10.1016/j.bandl.2014.09.005 [PubMed: 25306506]
- Krishnan A, Gandour JT, Suresh CH 2015a Experience-dependent enhancement of pitch-specific responses in the auditory cortex is limited to acceleration rates in normal voice range. *Neuroscience* 303, 433–45. 10.1016/j.neuroscience.2015.07.015 [PubMed: 26166727]
- Krishnan A, Gandour JT, Suresh CH 2015b Pitch processing of dynamic lexical tones in the auditory cortex is influenced by sensory and extrasensory processes. *Eur. J. Neurosci* 41, 1496–1504. 10.1111/ejn.12903 [PubMed: 25943576]
- Krishnan A, Gandour JT, Suresh CH 2016 Language-experience plasticity in neural representation of changes in pitch salience. *Brain Res* 1637, 102–117. 10.1016/j.brainres.2016.02.021 [PubMed: 26903418]

- Krishnan A, Suresh CH, Gandour JT 2017a Differential sensitivity to changes in pitch acceleration in the auditory brainstem and cortex. *Brain Lang* 169, 22–27. 10.1016/j.bandl.2017.01.014 [PubMed: 28237533]
- Krishnan A, Suresh CH, Gandour JT 2017b Changes in pitch height elicit both language-universal and language-dependent changes in neural representation of pitch in the brainstem and auditory cortex. *Neuroscience* 346, 52–63. 10.1016/j.neuroscience.2017.01.013 [PubMed: 28108254]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V 2014b Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia* 59, 1–12. 10.1016/j.neuropsychologia.2014.04.006 [PubMed: 24751993]
- Krishnan A, Gandour JT, Ananthakrishnan S, Vijayaraghavan V 2015c Language experience enhances early cortical pitch-dependent responses. *J. Neurolinguistics* 33, 128–148. 10.1016/j.jneuroling.2014.08.002 [PubMed: 25506127]
- Krishnan A, Gandour JT, Xu Y, Suresh CH 2017c Language-dependent changes in pitch-relevant neural activity in the auditory cortex reflect differential weighting of temporal attributes of pitch contours. *J. Neurolinguistics* 41, 38–49. [PubMed: 28713201]
- Krishnan A, Bidelman GM, Smalt CJ, Ananthakrishnan S, Gandour JT 2012b Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia* 50, 2849–59. 10.1016/j.neuropsychologia.2012.08.013 [PubMed: 22940428]
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lutkenhoner B 2003 Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb. Cortex* 13, 765–72. [PubMed: 12816892]
- Kwok VPY, Dan G, Yakpo K, Matthews S, Fox PT, Li P, Tan LH 2017 A Meta-Analytic Study of the Neural Systems for Auditory Processing of Lexical Tones. *Front. Hum. Neurosci* 11 10.3389/fnhum.2017.00375
- Li P, Sepanski S, Zhao X 2006 Language history questionnaire: A web-based interface for bilingual research. *Behav. Res. Methods* 38, 202–210. [PubMed: 16956095]
- Nabelek AK, Dagenais PA 1986a Vowel errors in noise and in reverberation by hearing-impaired listeners. *J. Acoust. Soc. Am* 80, 741–8. [PubMed: 3760327]
- Nabelek AK, Dagenais PA 1986b Vowel errors in noise and in reverberation by hearing-impaired listeners. *J. Acoust. Soc. Am* 80, 741–8. [PubMed: 3760327]
- Nabelek AK, Letowski TR 1988 Similarities of vowels in nonreverberant and reverberant fields. *J. Acoust. Soc. Am* 83, 1891–9. [PubMed: 3403805]
- Nabelek AK, Letowski TR, Tucker FM 1989 Reverberant overlap- and self-masking in consonant identification. *J. Acoust. Soc. Am* 86, 1259–65. [PubMed: 2808901]
- Okamoto H, Stracke H, Bermudez P, Pantev C 2011 Sound processing hierarchy within human auditory cortex. *J. Cogn. Neurosci* 23, 1855–63. 10.1162/jocn.2010.21521 [PubMed: 20521859]
- Oldfield RC 1971 The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97–113. [PubMed: 5146491]
- Oxenham AJ 2012 Pitch perception. *J. Neurosci* 32, 13335–8. 10.1523/JNEUROSCI.3815-12.2012 [PubMed: 23015422]
- Pantev C, Roberts LE, Schulz M, Engelen A, Ross B 2001 Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12, 169–74. [PubMed: 11201080]
- Parbery-Clark A, Skoe E, Kraus N 2009 Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci* 29, 14100–7. 10.1523/JNEUROSCI.3256-09.2009 [PubMed: 19906958]
- Penagos H, Melcher JR, Oxenham AJ 2004 A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *J. Neurosci* 24, 6810–5. 10.1523/JNEUROSCI.0383-04.2004 24/30/6810 [pii] [PubMed: 15282286]
- Plack CJ, Oxenham AJ, Fay RR, (Eds.) 2005 Pitch: neural coding and perception, Vol. 24 Springer, New York.
- Plomp R, Steeneken HJM 1973 Place dependence of timbre in reverberant sound fields. *Acustica* 28, 50–59.
- Rammsayer TH, Altenmüller E 2006 Temporal information processing in musicians and nonmusicians. *Music Perception* 24, 37–48.

- Sabine WC 1922 Collected papers on acoustics Harvard University Press, Cambridge, MA.
- 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. 10.1002/(SICI)1096-9861(19960715)371:1<15::AID-CNE2>3.0.CO;2-O [pii] 10.1002/(SICI)1096-9861(19960715)371:1<15::AID-CNE2>3.0.CO;2-O [PubMed: 8835717]
- Sayles M, Winter IM 2008 Reverberation challenges the temporal representation of the pitch of complex sounds. *Neuron* 58, 789–801. S0896-6273(08)00303-6 [pii] 10.1016/j.neuron.2008.03.029 [PubMed: 18549789]
- Sayles M, Stasiak A, Winter IM 2015 Reverberation impairs brainstem temporal representations of voiced vowel sounds: challenging “periodicity-tagged” segregation of competing speech in rooms. *Front. Syst. Neurosci* 8, 1–19.
- Schonwiesner M, Zatorre RJ 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. 10.1007/s00221-008-1286-z [PubMed: 18236034]
- Skoe E, Krizman J, Spitzer E, Kraus N 2015 Prior experience biases subcortical sensitivity to sound patterns. *J. Cogn. Neurosci* 27, 124–40. 10.1162/jocn_a_00691 [PubMed: 25061926]
- Somers M, Aukes MF, Ophoff RA, Boks MP, Flier W, de Visser KC, Kahn RS, Sommer IE 2015 On the relationship between degree of hand-preference and degree of language lateralization. *Brain Lang* 144, 10–5. 10.1016/j.bandl.2015.03.006 [PubMed: 25880901]
- Song JH, Skoe E, Wong PCM, Kraus N 2008 Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cogn. Neurosci* 20, 1892–1902. [PubMed: 18370594]
- Suga N, Gao E, Zhang Y, Ma X, Olsen JF 2000 The corticofugal system for hearing: recent progress. *Proc. Natl. Acad. Sci. U. S. A* 97, 11807–14. [PubMed: 11050213]
- Suga N, Ma X, Gao E, Sakai M, Chowdhury SA 2003 Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. *Speech Communication* 41, 189–200.
- Suresh CH, Krishnan A, Gandour JT 2017 Language experience-dependent advantage in pitch representation in the auditory cortex is limited to favorable signal-to-noise ratios. *Hear. Res* 355, 42–53. 10.1016/j.heares.2017.09.006 [PubMed: 28927640]
- Wang DL, Brown GJ 2006 Computational Auditory Scene Analysis Wiley, Hoboken, NJ.
- Watkins AJ 2005 Perceptual compensation for effects of reverberation in speech identification. *J. Acoust. Soc. Am* 118, 249–62. [PubMed: 16119347]
- Wong PC, Perrachione TK 2007 Learning pitch patterns in lexical identification by native English-speaking adults. *Appl. Psycholinguist* 28, 565–585.
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N 2007 Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci* 10, 420–2. nn1872 [pii] 10.1038/nn1872 [PubMed: 17351633]
- Xu Y 1997 Contextual tonal variations in Mandarin. *J. Phonetics* 25, 61–83.
- Zatorre RJ, Gandour JT 2008 Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci* 363, 1087–104. J412P80575385013 [pii] 10.1098/rstb.2007.2161 [PubMed: 17890188]
- Zatorre RJ, Baum SR 2012 Musical melody and speech intonation: Singing a different tune. *PLoS Biol* 10, e1001372 10.1371/journal.pbio.1001372 [PubMed: 22859909]
- Zatorre RJ, Belin P, Penhune VB 2002 Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci* 6, 37–46. [PubMed: 11849614]

- Experience-dependent FFR and CPR responses are maintained under reverberation
- Rightward asymmetry of CPR is maintained in adverse listening conditions
- Degradation in pitch representation at both levels of brain is language independent
- Differential effects of FFR and CPR reflect stronger top-down pitch processing

Figure 1.

Stimulus paradigm consists of two segments (A) 250 ms pitch segment (red) preceded by 800 ms noise segment (black). Pitch segment shows robust periodicity in the Dry stimulus condition. Waveform (B) and spectrogram (C) of the pitch segment are shown for each stimulus. Both reveal degradation of waveform periodicity and blurring of curvilinear pitch representation with increasing reverberation from Dry to Moderate. F0 contour (white) is superimposed in the spectrogram of the Dry stimulus. See also caption to Figure 5.

Figure 2.

Grand average waveforms of CPR responses obtained for the Chinese and English groups for each reverberant condition at the Fz electrode site. Na, Pb, and Nb are the most robust, pitch-relevant CPR components (highlighted in gray). CPR waveforms show that amplitude of Na, Pb, and Nb appear to be larger for the Chinese group as compared to the English when evoked by pitch stimuli across conditions. Solid black horizontal bar indicates the duration of each stimulus. CPR, cortical pitch response. To interpret mention of colors, the reader is referred to the online version of this article.

Figure 3. Mean peak-to-peak amplitude of CPR Na-Pb elicited by each of the four reverberant conditions at the Fz electrode site. Na-Pb amplitude for the Chinese group is larger compared to the English group regardless of condition. Error bars = ± 1 SE.

Figure 4.

Mean laterality indices ($T8-T7/T8+T7$) of CPR Na-Pb elicited by the four reverberant conditions. A rightward asymmetry ($T8 > T7$) is observed in the Chinese group across conditions. No significant asymmetries in Na-Pb amplitude are observed in the English group. Error bars = ± 1 SE.

Figure 5.

FFR waveforms (left) and spectrograms (right) evoked by Dry and reverberant (Slight, Mild, Moderate) stimulus conditions. Chinese FFR waveforms (red) are overlaid on English (blue). Chinese show larger amplitudes than English listeners particularly during the dynamic portion of the response (160–240 ms). Regardless of group, response amplitude appears to decrease with increasing reverberation. Solid white and red horizontal bars are superimposed in the spectrogram of the Moderate stimulus to designate the location and duration of FFR slow- (white) and fast-acceleration (red) segments, respectively. FFR, frequency following response; F0, voice fundamental frequency.

Figure 6.

Mean F0 amplitude of FFR slow- (top panel) and fast-acceleration (bottom panel) segments evoked by Dry and reverberant (Slight, Mild, Moderate) stimulus conditions. F0 magnitude of the Chinese group was larger than English irrespective of acceleration rate. Error bars = ± 1 SE.

Figure 7.

Mean normalized, magnitude of changes in amplitude in reference to the Dry condition for CPR Na–Pb (top panel) and FFR slow- (middle panel) and fast-acceleration (bottom panel) pitch segments. Pooled across language groups, change in normalized magnitude from the Dry-Slight condition was smaller than Dry-Mild and Dry-Moderate; Dry-Mild, in turn, was smaller than Dry-Moderate. At both cortical and brainstem levels of representation, response amplitude change gets larger with increasing reverberation. For CPR Na–Pb, the amount of change in magnitude of Dry-Mild and Dry-Moderate conditions is greater than Dry-Slight. For both FFR pitch segments, the amount of change in magnitude of Dry-Mild and Dry-Moderate is greater than Dry-Slight; Dry-Moderate is also larger than Dry-Mild. These data

imply that the brainstem response (FFR) is more susceptible to reverberant effects on pitch processing than the cortical response (Na–Pb).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript