

## Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans

Ananthanarayan Krishnan<sup>a,\*</sup>, Gavin M. Bidelman<sup>b</sup>, Christopher J. Smalt<sup>c</sup>, Saradha Ananthakrishnan<sup>a</sup>, Jackson T. Gandour<sup>a</sup>

<sup>a</sup> Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907-2038, USA

<sup>b</sup> Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, ON, Canada

<sup>c</sup> School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-2038, USA

### ARTICLE INFO

#### Article history:

Received 11 June 2012

Received in revised form

1 August 2012

Accepted 15 August 2012

Available online 23 August 2012

#### Keywords:

Fundamental frequency

Pitch encoding

Auditory brainstem

Frequency-following response

Auditory cortex

Cortical pitch response

Pitch discrimination

### ABSTRACT

Neural representation of pitch-relevant information at both the brainstem and cortical levels of processing is influenced by language or music experience. However, the functional roles of brainstem and cortical neural mechanisms in the hierarchical network for language processing, and how they drive and maintain experience-dependent reorganization are not known. In an effort to evaluate the possible interplay between these two levels of pitch processing, we introduce a novel electrophysiological approach to evaluate pitch-relevant neural activity at the brainstem and auditory cortex concurrently. Brainstem frequency-following responses and cortical pitch responses were recorded from participants in response to iterated rippled noise stimuli that varied in stimulus periodicity (pitch salience). A control condition using iterated rippled noise devoid of pitch was employed to ensure pitch specificity of the cortical pitch response. Neural data were compared with behavioral pitch discrimination thresholds. Results showed that magnitudes of neural responses increase systematically and that behavioral pitch discrimination improves with increasing stimulus periodicity, indicating more robust encoding for salient pitch. Absence of cortical pitch response in the control condition confirms that the cortical pitch response is specific to pitch. Behavioral pitch discrimination was better predicted by brainstem and cortical responses together as compared to each separately. The close correspondence between neural and behavioral data suggest that neural correlates of pitch salience that emerge in early, preattentive stages of processing in the brainstem may drive and maintain with high fidelity the early cortical representations of pitch. These neural representations together contain adequate information for the development of perceptual pitch salience.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

Voice pitch, an important information-bearing perceptual component of language and music, provides an excellent window for studying experience-dependent effects on both brainstem and cortical components of a well-coordinated, hierarchical processing network. There is growing empirical evidence to support the notion that the neural representation of pitch-relevant information at both brainstem and cortical levels of processing is influenced by one's experience with language and/or music (Kraus & Banai, 2007; Krishnan & Gandour, 2009; Patel & Iversen, 2007). It is well known from animal studies that neural processes mediating experience-dependent plasticity for pitch at

the brainstem and cortical levels may be well-coordinated based on neuroanatomical evidence of ascending and descending pathways (Huffman & Henson, 1990; Kelly & Wong, 1981; Saldana, Feliciano, & Mugnaini, 1996) and physiological evidence of improved signal representation in subcortical structures mediated by the corticofugal system (Suga, Ma, Gao, Sakai, & Chowdhury, 2003; Yan & Suga, 1998; Zhou & Jen, 2000). Human electrophysiological studies have also shown enhanced brainstem neural activity in individuals with short-term auditory training (Russo, Nicol, Zecker, Hayes, & Kraus, 2005; Song, Skoe, Wong, & Kraus, 2008), long-term linguistic experience (Krishnan & Gandour, 2009; Krishnan, Gandour, & Bidelman, 2012; Krishnan, Xu, Gandour, & Cariani, 2005), and musical training (Bidelman, Gandour, & Krishnan, 2011a; Bidelman & Krishnan, 2009; Musacchia, Sams, Skoe, & Kraus, 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007). Results showing correlation between brainstem and cortical responses in musicians suggest that brainstem neural representations of pitch, timing and timbre cues and cortical

\* Corresponding author. Tel.: +1 765 494 3793; fax: +1 765 494 0771.

E-mail addresses: [rkrish@purdue.edu](mailto:rkrish@purdue.edu) (A. Krishnan), [gbidelman@rotman-baycrest.on.ca](mailto:gbidelman@rotman-baycrest.on.ca) (G.M. Bidelman), [csmalt@purdue.edu](mailto:csmalt@purdue.edu) (C.J. Smalt), [sanantha@purdue.edu](mailto:sanantha@purdue.edu) (S. Ananthakrishnan), [gandour@purdue.edu](mailto:gandour@purdue.edu) (J.T. Gandour).

response timing are shaped in a coordinated manner through corticofugal modulation of subcortical afferent circuitry (Musacchia, Strait, & Kraus, 2008). However, little is known about how language (music) experience shapes pitch at each level of the processing hierarchy or how it modulates the nature of the interplay between them. The scalp-recorded brainstem frequency following response (FFR) and the cortical pitch onset response (POR) representing neural activity relevant to pitch at brainstem and cortical levels, respectively, provide a physiologic window to evaluate the hierarchical organization of pitch processing along the auditory pathway.

The FFR reflects sustained phase-locked activity in a population of neural elements within the rostral brainstem (see Chandrasekaran & Kraus, 2010; Krishnan, 2007, for reviews). It has provided numerous insights into pitch encoding of ecologically-relevant stimuli including speech (Krishnan & Gandour, 2009) and music (Bidelman, Krishnan, & Gandour, 2011). Furthermore, the FFR has revealed that experience-dependent plasticity enhances neural representation of pitch in native speakers of a tone language (Krishnan, Swaminathan, & Gandour, 2009; Krishnan et al., 2005) and individuals with extensive music experience (Bidelman, Gandour, & Krishnan, 2011b; Lee, Skoe, Kraus, & Ashley, 2009; Wong et al., 2007). Finally, pitch-relevant information preserved in the FFR is strongly correlated with perceptual pitch measures (Bidelman & Krishnan, 2011; Krishnan, Bidelman, & Gandour, 2010; Krishnan & Plack, 2009; Parbery-Clark, Skoe, Lam, & Kraus, 2009) suggesting that acoustic features relevant to pitch are already emerging in representations at the level of the brainstem.

The POR, as recorded using magnetoencephalography (MEG), is thought to reflect synchronized cortical neural activity specific to pitch (Chait, Poeppel, & Simon, 2006; Krumbholz, Patterson, Seither-Preisler, Lammertmann, & Lutkenhoner, 2003; Seither-Preisler, Patterson, Krumbholz, Seither, & Lutkenhoner, 2006). For example, POR latency and magnitude has been shown to depend on specific features of pitch (e.g., salience, fundamental frequency). A more robust POR with shorter latency is observed for stimuli with stronger pitch salience compared to ones with weaker pitch salience. In order to disentangle the POR from the obligatory onset responses (P1–N1–P2), Krumbholz et al. (2003) utilized a novel stimulus paradigm in which a continuous sound is constructed using an initial segment of noise with no pitch (that evokes only the onset components), followed by a pitch-eliciting segment of iterated rippled noise (IRN) matched in intensity and overall spectral profile. Interestingly, the POR is evoked only for this noise-to-pitch transition and not for the pitch-to-noise stimulus transition. Source analyses (Gutschalk, Patterson, Rupp, Uppenkamp, & Scherg, 2002; Gutschalk, Patterson, Scherg, Uppenkamp, & Rupp, 2004; Krumbholz et al., 2003), corroborated by human depth electrode recordings (cf. Griffiths et al., 2010; Schonwiesner & Zatorre, 2008), indicate that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing (Bendor & Wang, 2005; Griffiths, Buchel, Frackowiak, & Patterson, 1998; Johnsrude, Penhune, & Zatorre, 2000; Penagos, Melcher, & Oxenham, 2004; Zatorre, 1988). Given both its sensitivity and consistency across a number of studies, the POR offers an excellent window for studying early cortical representations of pitch. Our preliminary POR data, extracted from scalp-recorded EEG, yielded multiple peaks in addition to pitch onset. We therefore have chosen to designate this scalp-recorded neural activity as cortical pitch response (CPR).

Recently, Krishnan, Gandour, Smalt, and Bidelman (2010) demonstrated that the degree of neural periodicity (i.e., pitch-relevant phase-locked neural activity) as reflected in the brainstem FFR accurately predicts the perceptual salience of IRN pitch. Moreover, strong correlations were observed between neural and behavioral measures of pitch. These findings support the notion that early sensory level representations of pitch relevant

information in the brainstem may play an important role in formulating and/or shaping pitch percepts (Baumann et al., 2011; Langner, 1983; Pantev, Hoke, Lutkenhoner, & Lehnertz, 1989; Zatorre, Evans, & Meyer, 1994). While previous research has documented details of pitch encoding mechanisms at brainstem, cortical, and perceptual levels of processing separately, we are not cognizant of any published accounts that examine the interplay and coordination across subcortical and cortical levels of processing.

In an effort to increase our understanding of the organization of the hierarchical network underlying pitch processing and the nature of the interplay between levels of processing along the hierarchy, we introduce herein a novel experimental approach whereby neural representation of pitch-relevant information at brainstem (FFR) and cortical (CPR) levels can be recorded simultaneously in response to IRN stimuli varying in pitch salience. We further compare these neural indices to perceptual measures of pitch salience. This combined approach gives us a unique window to examine the coordination between different levels of pitch processing in real time, which may otherwise be obscured by inferences drawn from separate evaluation of neural responses evoked by different stimulation/acquisition paradigms or comparisons across studies.

## 2. Materials and methods

### 2.1. Participants

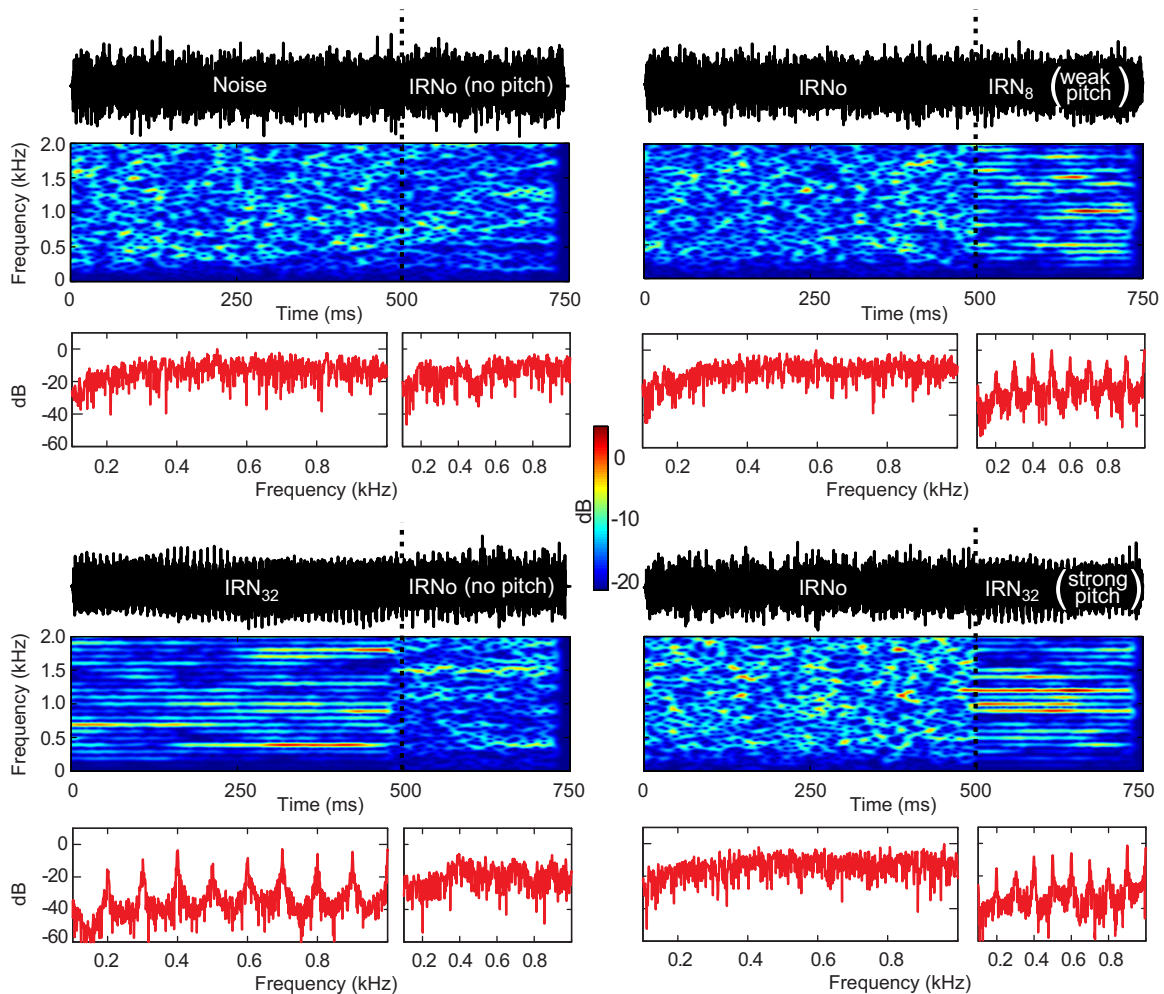
Thirteen Purdue University students (five male, eight female) were recruited to participate in the experiment. All exhibited normal hearing sensitivity at octave frequencies between 500 and 4000 Hz and reported no previous history of neurological or psychiatric illnesses. Participants were closely matched in age ( $23.9 \pm 3.1$  years), years of formal education ( $17.9 \pm 2.2$  years), and were strongly right handed ( $90.4 \pm 15.6\%$ ) as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). Two participants had more than 10 years of instrumental musical training; all others, less than three years. Participants were paid and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

### 2.2. Stimuli

Four IRN stimuli were employed, each consisting of two segments: an initial 500 ms segment followed by a 250 ms segment (Fig. 1). For each stimulus, the two segments were crossfaded with 5 ms cos<sup>2</sup> ramps to produce the following four transitions: Noise eac (read “no-pitch IRN” IRNitch IRNitch IRN, the two segments were crossfaded with 5 ms cos<sup>2</sup> ramps, IRN0, IRN8, IRN32, and pIRN0, respectively). The overall RMS level of each segment was equated such that there was no discernible difference in intensity between initial and final segments. Temporal and spectral characteristics of the stimuli are shown in Fig. 1.

IRN is a complex pitch-evoking stimulus which has been widely used for examining temporal pitch mechanisms and pitch salience, as it allows one to systematically manipulate the temporal periodicity and hence pitch salience of a stimulus. Yet, IRN lacks the prominent temporal envelope typical of most signals carrying pitch. Studies show that the pitch of IRN corresponds to the reciprocal of the delay ( $1/d$ ) and that its salience grows with the number of iterations (Krishnan et al., 2010; Patterson, Handel, Yost, & Datta, 1996; Yost, 1996a; Yost, 1996b). IRN stimuli were created by delaying Gaussian noise (80–4000 Hz) and adding it back on itself in a recursive manner, producing a pitch percept corresponding to the reciprocal of the delay ( $d$ ) (Yost, 1996a). To examine the effects of changing pitch salience on the FFR, CPR, and behavioral measures of pitch, two different iterations steps were used to create the sensation of either a weak ( $n=8$ ) or a strong ( $n=32$ ) steady state pitch corresponding to 100 Hz ( $d=10$  ms).

In addition to the Noise control stimulus (matched in bandwidth to experimental stimuli), IRN0 served as a second control stimulus. Informal listening to IRN0 by trained musicians confirmed that it does not support the production/identification of musical melody. It therefore does not satisfy the most conservative definition of pitch (ASA, 1960; Plack, 2005, p. 2). Like noise, it should not evoke an electrophysiologic response specific to pitch. IRN0 was created by manipulating an IRN32 segment by moving a sliding window (length equal to  $d$ ) across its temporal waveform and randomizing the phase within each window, thereby removing fine temporal structure (Barker, Plack, & Hall, 2012). While this manipulation removes the sensation of pitch, it retains the broad spectro-temporal features germane to IRN. Here, utilizing IRN0 removes any concomitant



**Fig. 1.** Waveform temporal and spectral characteristics of the four stimulus conditions. Vertical dotted lines (black) in each panel demarcate the transition from the initial no pitch segment (500 ms) to the final pitch segment (250 ms) to illustrate the experimental paradigm used to acquire brainstem and cortical responses concurrently. Individual FFTs corresponding to the spectra computed within each of the two stimulus segments are shown in red. Panels in the left column show the control stimulus conditions in which transitions were from either a broad-band noise (top) or salient pitch (bottom) to IRN0 (i.e., no pitch). In contrast, the two panels in the right column show stimulus conditions in which transitions were from IRN0 to either a weak (IRN8, top) or a strong (IRN32, bottom) pitch. Note the clear bands of energy corresponding to the harmonics of the 100 Hz fundamental frequency in both the spectrograms and FFT spectra for the pitch-eliciting stimuli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acoustic confounds allowing us to confirm that the CPR is an electrophysiologic response specific to the pitch of IRN rather than to one of its other latent acoustic features, e.g., spectro-temporal flux (cf. Barker et al., 2012; Hall & Plack, 2009). IRN0 also serves as a control to confirm that the CPR is elicited by pitch onset rather than pitch offset.

### 2.3. Brainstem and cortical evoked response data acquisition

Participants reclined comfortably in an electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts), ignore the sounds they heard, and were allowed to sleep throughout the duration of the recording procedure (~75% fell asleep). The continuous EEG was recorded from each participant in response to binaural presentation at 80 dB SPL through magnetically shielded insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA). Stimuli were presented with fixed (rarefaction) onset polarity at a repetition rate of 0.93/s. Stimulus presentation order was randomized both within and across participants. Control of the experimental protocol was accomplished by a signal generation and data acquisition system (Intelligent Hearing Systems; Miami, FL, USA) using a sampling rate of 10 kHz.

The continuous EEG was recorded differentially between Ag and AgCl scalp electrodes placed on the midline of the forehead at the hairline (~Fpz) referenced to the linked mastoids (A2/A1). Another electrode placed on the mid-forehead served as the common ground. A vertical electrode montage was selected since it provides the optimal configuration for recording brainstem responses (Galbraith et al., 2000). Inter-electrode impedances were maintained  $\leq 1$  k $\Omega$ , amplified by 100,000, and low-pass filtered online below 5000 Hz (6 dB/octave roll-off). Raw

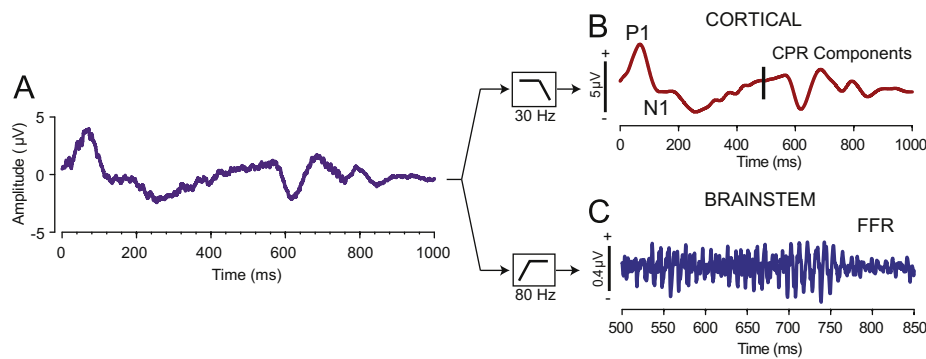
EEGs were then divided into epochs using an analysis time window from 0 to 1075 ms (0 ms = stimulus onset). Evoked responses were obtained by time-domain averaging each epoch over the duration of the EEG recording. A total of 1000 averages were collected for each condition in blocks of 250 sweeps. Sweeps containing activity exceeding  $\pm 75$   $\mu$ V were rejected as artifacts and excluded from the final average. The experimental protocol took about 2 h to complete.

### 2.4. Analyses of brainstem FFR and cortical CPR

Brainstem and cortical responses were extracted from the same sweep-epoch EEG through offline zero-phase FIR filtering (Fig. 2). Waveforms were high-pass (80–1500 Hz) or low-pass (1–30 Hz) filtered offline to emphasize brainstem (FFR) or cortical (CPR) activity, respectively (Musacchia et al., 2008).

#### 2.4.1. Brainstem FFR neural pitch salience measured using harmonic sieve analysis

FFTs were first computed from each FFR to index the magnitude of spectral information contained in each brainstem response. Neural pitch salience was then estimated from each spectra using harmonic templates whereby a series of sieves selected spectral activity at a given frequency and its integer multiples (Bidelman & Heinz, 2011; Bidelman & Krishnan, 2009, 2011; Cedolin & Delgutte, 2005). Each sieve template (representing a single pitch) was composed of 5 Hz wide bins situated at a fundamental frequency ( $f_0$ ) and its integer multiples (i.e.,  $2f_0$ ,  $3f_0$ ,  $4f_0$ ,  $nf_0$ ; for  $nf_0 < 1000$  Hz). This analysis essentially implements the classic pattern recognition model of pitch in which a central processor matches harmonic information contained in the neural response to an internal template in order to compute the heard pitch (e.g., Goldstein, 1973; Terhardt, Stoll, & Seewann, 1982). All sieve templates with  $f_0$ s between 25 and 1000 Hz (5 Hz steps) were considered. For each template, the degree



**Fig. 2.** Illustration of the derivation of cortical and brainstem response from the EEG. For each participant, the raw EEG was recorded with online filters set between 0 and 5000 Hz, epoched, and averaged (panel A). This averaged waveform was then either low-pass filtered (1–30 Hz) offline to derive the cortical CPR components (Panel B) or band-pass filtered (80–1500 Hz) to derive the brainstem FFR (panel C). The vertical line at 500 ms marks the onset of the pitch eliciting segment of the stimulus.

of salience was estimated by dividing the mean density of activity falling within the sieve bins by the mean density of activity in the whole FFT, thus providing a contrast between pitch-related activity and the background energy (unrelated to pitch). Activity falling within sieve windows adds to the total estimated pitch salience while information falling outside the windows reduces the total pitch salience (Cariani & Delgutte, 1996; Cedolin & Delgutte, 2005). The output from each template was then concatenated as a function of the template's  $f_0$  to construct a running neural pitch salience curve representing the relative strength of all possible pitches present in the FFR response. The salience magnitude corresponding to a template of 100 Hz was taken as a singular measure of FFR pitch salience for a given stimulus condition. The noise floor was estimated by averaging voltage values over a 40 Hz band on either side of the salience magnitude peak for the 100 Hz template.

#### 2.4.2. Cortical CPR latency and magnitude

Obligatory components corresponding to the onset of energy in the stimulus (P1/N1) were measured from each cortical response. P1 was taken as the peak positivity within a search window between 40 and 100 ms; N1 as the maximum negativity between 85 and 200 ms. Responses corresponding to the onset of pitch were taken as a measure of CPR magnitude, computed as the voltage difference between the maximum negative deflection within the 600–650 ms time window and the following positive component. The time at which the maximum negativity occurred was taken as a measure of CPR latency (~630 ms). The noise floor was estimated by obtaining an average voltage over a 100 ms time window preceding the onset of the pitch-eliciting segment at 500 ms.

#### 2.5. Behavioral estimates of pitch salience

We adopted a fundamental-frequency difference limen ( $F_0$  DL) paradigm as an objective index to estimate the perceptual pitch salience of the IRN stimuli. The rationale behind using an  $F_0$  DL is that pitch discrimination should improve with increasing salience of pitch, i.e., with increasing IRN iteration steps (Hall & Plack, 2009; Krishnan et al., 2010). Testing consisted of three conditions (IRNO, IRN8, IRN32) in which participants performed a two-alternative forced choice task (2-AFC). For a given trial within a condition, they heard two sequential intervals, one containing a reference stimulus and one containing a comparison, assigned randomly. The reference pitch had a fixed  $F_0$  frequency of 100 Hz; that of the comparison was always greater (i.e., higher  $F_0$ ). The task was to identify the interval which contained a higher sounding pitch. Discrimination thresholds were measured using a two-down, one-up adaptive rule tracking 71% performance (Levitt, 1971). Following two consecutive correct responses, the frequency difference of the comparison was decreased for the subsequent trial, and increased following a single incorrect response. Sixteen reversals were measured and the geometric mean of the last 12 taken as the individual's  $F_0$  DL, that is, the minimum frequency difference needed to detect a change in pitch for a given stimulus condition.

#### 2.6. Statistical analysis

Separate one-way, mixed-model ANOVAs were conducted on FFR pitch salience; magnitude and latency of P1, N1; and magnitude and latency of CPR in order to evaluate the effects of pitch salience on the neural representation of pitch-relevant information at the brainstem and cortical levels. Subjects were treated as a random factor, stimuli as a between-subjects factor. Both CPR magnitude and FFR pitch salience were square-root transformed, after taking their modulus, to satisfy normality and homogeneity of variance assumptions necessary for a parametric ANOVA. P1/N1 magnitudes and latency measures were left untransformed.

Behavioral  $F_0$  DL scores were first log-transformed before being submitted to an ANOVA (Demany & Semal, 2002; Micheyl, Delhommeau, Perrot, & Oxenham,

2006). Transformed  $F_0$  DLs were analyzed using a similar one-way model with stimulus as the between-subjects factor of interest (3 levels; IRNO, IRN8, IRN32). An a priori level of significance was set at  $\alpha=0.05$ . Tukey–Kramer corrections were used to adjust for multiple comparisons.

To evaluate the nature of the relationship between several levels of pitch processing, Pearson's correlations ( $r$ ) were computed between FFR neural pitch salience and cortical CPR magnitude, cortical CPR magnitude and behavioral  $F_0$  DLs, and FFR pitch salience and behavioral  $F_0$  DLs.

### 3. Results

#### 3.1. Brainstem FFR

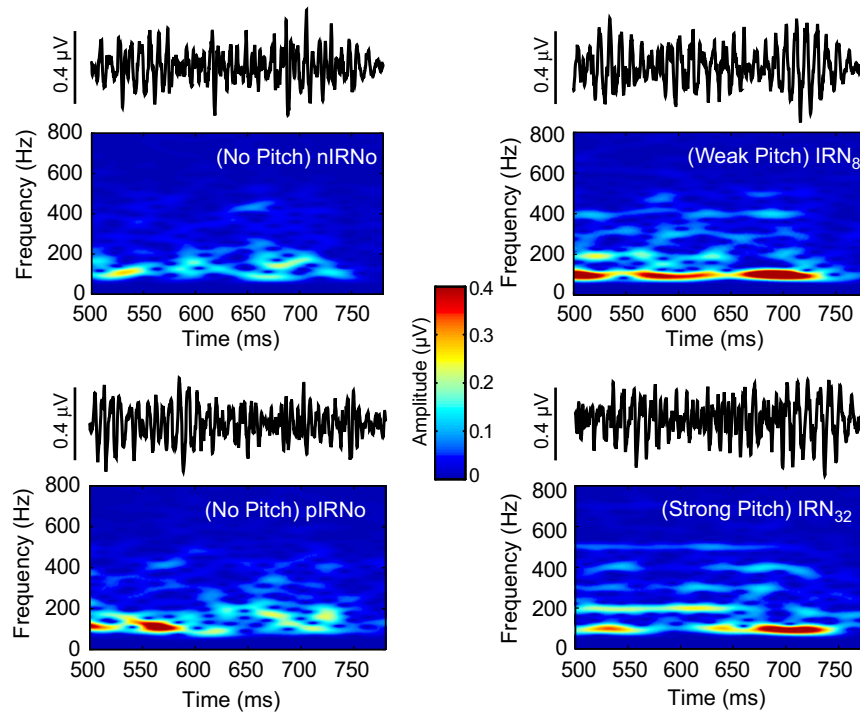
Derived FFR time-waveforms (500–780 ms) and their corresponding spectrograms in response to the four stimulus conditions are shown in Fig. 3. Qualitatively, time waveforms show clearer periodicity and larger amplitude with increasing iterations (i.e., IRN32 > IRN8 > IRNO ≈ nIRNO). Spectrograms, likewise, reveal clearer and more robust spectral bands at the harmonics of the 100 Hz fundamental frequency for stimuli evoking clear pitch (IRN8 and IRN32) as compared to IRNO conditions which do not evoke a recognizable pitch percept.

Consistent with these observations, mean FFR pitch salience is greater for the pitch-eliciting IRN8 and IRN32 stimulus conditions compared to the no-pitch control stimulus conditions, pIRNO and nIRNO (Fig. 4). Note that the pitch salience values of the control conditions approximate the noise floor and therefore do not reflect pitch-relevant, phase-locked neural activity. An omnibus ANOVA on FFR pitch salience revealed a significant effect of stimulus condition [ $F_{3,36}=53.33$ ,  $p<0.0001$ ]. Tukey–Kramer adjusted multiple comparisons indicated that FFR salience was more robust for IRN32 than for all other conditions. IRN8 also elicited greater salience than the two IRNO conditions. As expected, FFR salience evoked by the two control conditions did not differ from one another. This outcome confirms that these stimuli lack the necessary temporal regularity to evoke a measurable brainstem FFR. This pattern (IRN32 > IRN8 > pIRNO = nIRNO) indicates that brainstem activity becomes more robust with increasing temporal regularity of the stimulus such that stimuli of higher salience (IRN32) evoke more robust subcortical responses than stimuli with little (IRN8) or no salience (pIRNO and nIRNO).

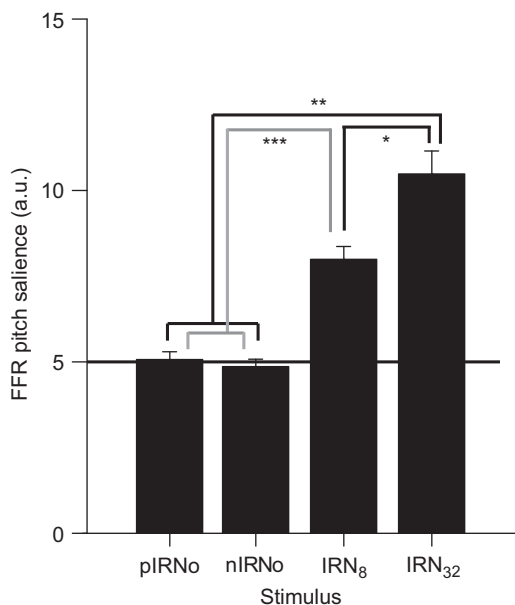
#### 3.2. Cortical response components

##### 3.2.1. Response morphology

Derived cortical response time-waveforms of the entire acquisition window are shown in Fig. 5. A clear P1/N1 onset complex is seen in the 50–150 ms time window which appears identical across all stimulus conditions. These obligatory components reflect the onset of

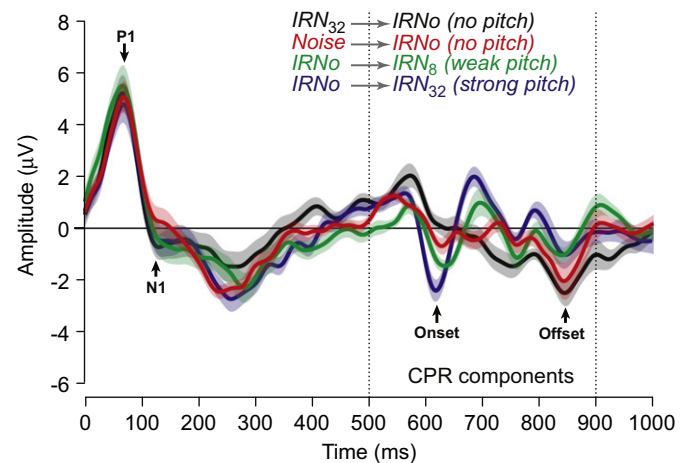


**Fig. 3.** Brainstem FFR waveforms and spectrograms as a function of stimulus temporal regularity. FFRs were extracted from evoked responses to the second segment following the transition from the initial precursor segment of the stimulus to the onset of pitch (IRN8, IRN32) or IRNO (500–780 ms time window). Note that pIRNO (pitch → IRNO) and nIRNO (Noise → IRNO) served as control conditions. Only responses to pitch-eliciting stimuli show clear sustained FFRs and their corresponding spectral components at harmonics of the  $F_0$  (right column). Observe the improved encoding from weak to strong pitch as indicated by the clearer pitch-relevant harmonics (i.e., spectral bands at integers of 100 Hz) of IRN32 relative to IRN8.



**Fig. 4.** Mean FFR pitch salience plotted (in arbitrary units) for the four stimulus conditions. The horizontal solid line across the conditions corresponds to the average noise floor. Note that responses to the control stimuli (pIRNO and nIRNO) are essentially indistinguishable from the noise floor. Asterisk(s) represent comparisons of neural pitch salience that are statistically significant. Error bars denote  $\pm 1$  SEM.

energy for each stimulus whose amplitudes and latencies are invariant to changes in pitch salience. The response to the pitch-eliciting stimulus segment is characterized by a series of successive biphasic components with no sustained negativity typically observed in the MEG-recorded POR. The CPR component is the prominent negativity around 630 ms. Systematic changes in CPR amplitude and



**Fig. 5.** Components of the cortical evoked response. Vertical dotted lines demarcate the final stimulus segment containing either pitch (IRN32, IRN8) or noise (IRNO) (see also Fig. 1). For all conditions, an obligatory P1/N1 complex is observed between 50 and 150 ms reflecting the responses to the onset of stimulus energy. A clear cortical pitch response (CPR) is observed as a prominent negative deflection ( $\sim 630$  ms) in the transition from noise to pitch (IRNO → IRN8, IRNO → IRN32). No response is observed in the transition from pitch to noise (IRN32 → IRNO) or from noise to noise (Noise → IRNO). CPR amplitude increases systematically with increasing stimulus temporal regularity.

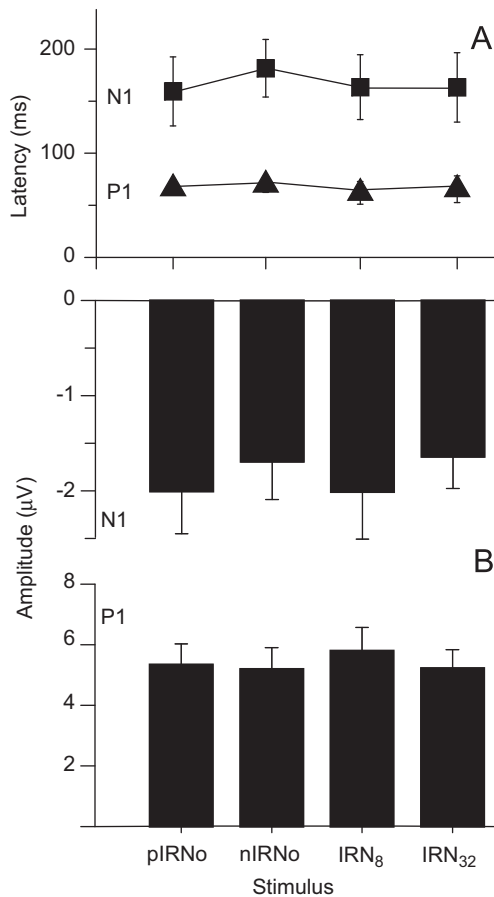
latency are observed across stimulus conditions. The prominence of the CPR complex increases with increasing regularity of the stimulus (IRN32 > IRN8). Weak or no CPR complex is discernible for IRNO no matter whether it is preceded by pitch (pIRNO) or noise (nIRNO), whereas a clear response is observed when transitioning from noise to pitch (IRN8; IRN32). Note that the smaller magnitude of the responses identified as CPR for the control conditions approximates the noise floor.

### 3.2.2. P1, N1, and CPR latency and magnitude

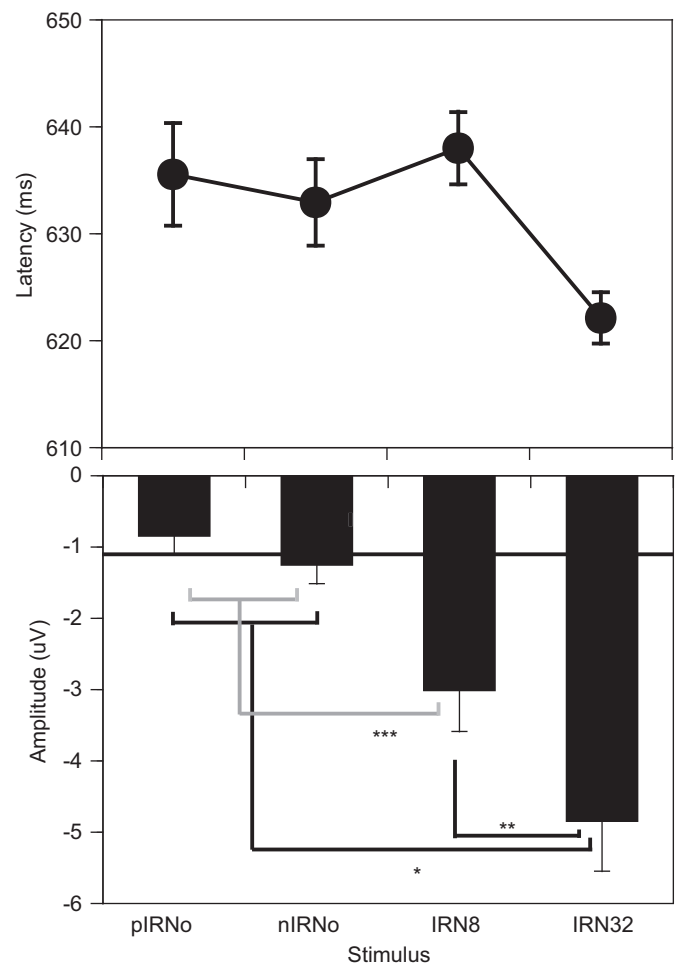
Mean latency (panel A) and magnitude for P1 and N1 (panel B) components are shown in Fig. 6. Separate ANOVAs indicated that both magnitude [P1:  $F_{3,36}=1.23$ ,  $p=0.315$ ; N1:  $F_{3,36}=0.52$ ,  $p=0.673$ ] and latency [P1:  $F_{3,36}=1.72$ ,  $p=0.1811$ ; N1:  $F_{3,36}=1.46$ ,  $p=0.2429$ ] of these components remained unaltered by manipulation of stimulus periodicity, meaning that both P1 and N1 amplitudes were invariant across conditions.

In contrast, increase in pitch salience produced shorter latency (Fig. 7, top panel) and greater amplitude (Fig. 7, bottom panel) for the CPR. ANOVA results showed a significant effect of stimulus temporal regularity on CPR latency [ $F_{3,36}=4.46$ ,  $p=0.0092$ ]. Post hoc Tukey–Kramer adjusted multiple comparisons indicated that the CPR evoked by the IRN32 condition occurred earlier ( $622.1 \pm 8.6$  ms) than that evoked by either IRN0 (nIRN0:  $632.8 \pm 17.2$  ms; pIRN0:  $635.6 \pm 17.4$  ms) or IRN8 ( $637.0 \pm 11.3$  ms) conditions (cf. Krumbholz et al., 2003). That is, a stimulus with stronger pitch salience (IRN32) evokes a CPR with shorter latency compared to a stimulus with weaker pitch salience (IRN8) or no discernible pitch (pIRN0; nIRN0).

Consistent with the latency data, an ANOVA on CPR amplitude revealed a significant effect of stimulus temporal regularity [ $F_{3,36}=26.15$ ,  $p < 0.0001$ ]. Tukey–Kramer comparisons indicated an identical pattern of contrasts for CPR magnitudes as found for FFR pitch salience (i.e.,  $IRN32 > IRN8 > pIRN0 = nIRN0$ ). That is, stimuli containing higher degrees of temporal regularity elicit stronger CPR magnitude than those containing poorer temporal regularity. These results indicate that latency and magnitude of the CPR is sensitive to the temporal regularity of IRN stimuli and therefore to changes in pitch salience.



**Fig. 6.** Mean latency (panel A), and amplitude (panel B) of cortical components P1, and N1 per stimulus condition. These obligatory onset components show invariance in both latency and amplitude across stimulus conditions. The invariance suggests that P1 and N1 are insensitive to changes in pitch salience.



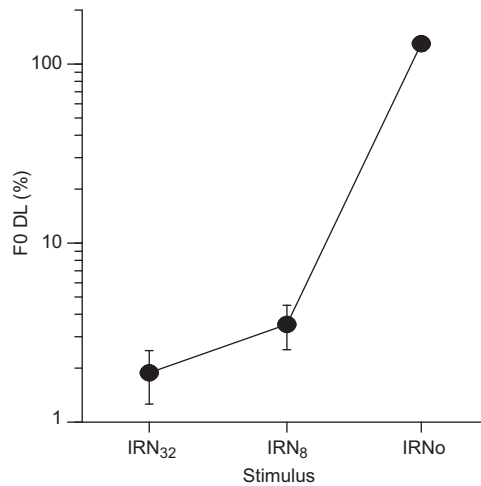
**Fig. 7.** Mean latency (top panel) and amplitude (bottom panel) of CPR for the four stimulus conditions. Both response latency and amplitude are altered by temporal regularity of the stimulus as an index to their sensitivity to pitch salience. The horizontal solid line ( $\sim -1 \mu\text{V}$ ) across the conditions corresponds to the average noise floor. Observe that CPR amplitudes for the control stimuli (pIRN0 and nIRN0) are essentially indistinguishable from the noise floor. Asterisk(s) represent neural pitch salience comparisons that are statistically significant. Error bar denotes  $\pm 1$  SEM.

### 3.3. Behavioral measures of pitch salience

Mean  $F_0$  DLs for the IRN32, IRN8, and IRN0 stimulus conditions show that the size of the  $F_0$  DL increases with decreasing temporal regularity of the stimulus (Fig. 8), meaning  $F_0$  discrimination ability gets poorer with decreasing pitch salience. An omnibus ANOVA on behavioral  $F_0$  DLs revealed a significant effect of stimulus [ $F_{2,24}=284.39$ ,  $p < 0.0001$ ]. Post-hoc comparisons indicated that listeners'  $F_0$  DLs were largest for IRN0 and decreased monotonically with an increasing number of iteration steps ( $n$ ). This pattern of performance ( $IRN32 < IRN8 < IRN0$ ) indicates that participants were better at detecting changes in pitch when listening to stimuli with higher pitch salience compared to when listening to stimuli with weaker pitch or no pitch.  $F_0$  DLs for the IRN0 condition were nearly two orders of magnitude larger (i.e., poorer) than in either the IRN8 or IRN32 condition.

### 3.4. Comparison of neural and behavioral representations of pitch relevant information

Regressions of neural (FFR pitch salience, CPR magnitude) and behavioral ( $F_0$  DL) measures against one another were conducted to determine the extent to which brainstem and/or cortical



**Fig. 8.** Mean  $F_0$  DLs for three stimulus conditions (IRN32, IRN8, IRN0) show a systematic increase in  $F_0$  DL thresholds with decreasing temporal regularity of the stimulus. The higher the  $F_0$  DL threshold, the poorer is the ability to discriminate changes in pitch salience. Error bar denotes  $\pm 1$  SEM.

representation of pitch-relevant information can predict perceptual responses to pitch (Fig. 9). The IRN32→IRN0 (pIRN0) condition was excluded from the correlation analyses because it was identical to Noise→IRN0 (nIRN0). Pearson's  $r$  computed between dependent variables indicated significant correlations between FFR neural pitch salience and cortical CPR magnitude (Fig. 9A:  $r_{\text{FFR,CPR}}=0.49$ ,  $p<0.0017$ ); cortical CPR magnitude and behavioral  $F_0$  DL (Fig. 9B:  $r_{\text{CPR,Behavior}}=-0.62$ ,  $p<0.0001$ ); and FFR salience and behavioral  $F_0$  DL (Fig. 9C:  $r_{\text{FFR,Behavior}}=-0.80$ ,  $p<0.0001$ ). We observe that in each case, stimuli devoid of pitch (IRN0) or eliciting only weak pitch (IRN8) cluster away from the stimulus that elicits a salient pitch (IRN32).

Fig. 9D shows a 3-dimensional scatter where each point in space represents a listener's brainstem, cortical, and behavioral responses in the  $x$ ,  $y$ , and  $z$  dimensions, respectively. Fitting a plane to the data revealed a significant correlation between all three measures ( $r_{\text{all}}=0.84$ ,  $p<0.0001$ , better than any single pairwise correlation alone (Fig. 9A–C). The improvement in the ability to predict  $F_0$  DLs using both neural measures was supported by a step-wise multiple regression analysis (CPR: adjusted  $R^2=0.37$ ; FFR:  $R^2=0.63$ ; FFR+CPR:  $R^2=0.70$ ). This implies that weaker representation of pitch in the brainstem and auditory cortex is associated with poorer perceptual performance (i.e., larger DLs; upper portion of the plane), whereas stronger neural representation at both levels is associated with improved performance (i.e., lower DLs; lower portion of the plane).

#### 4. Discussion

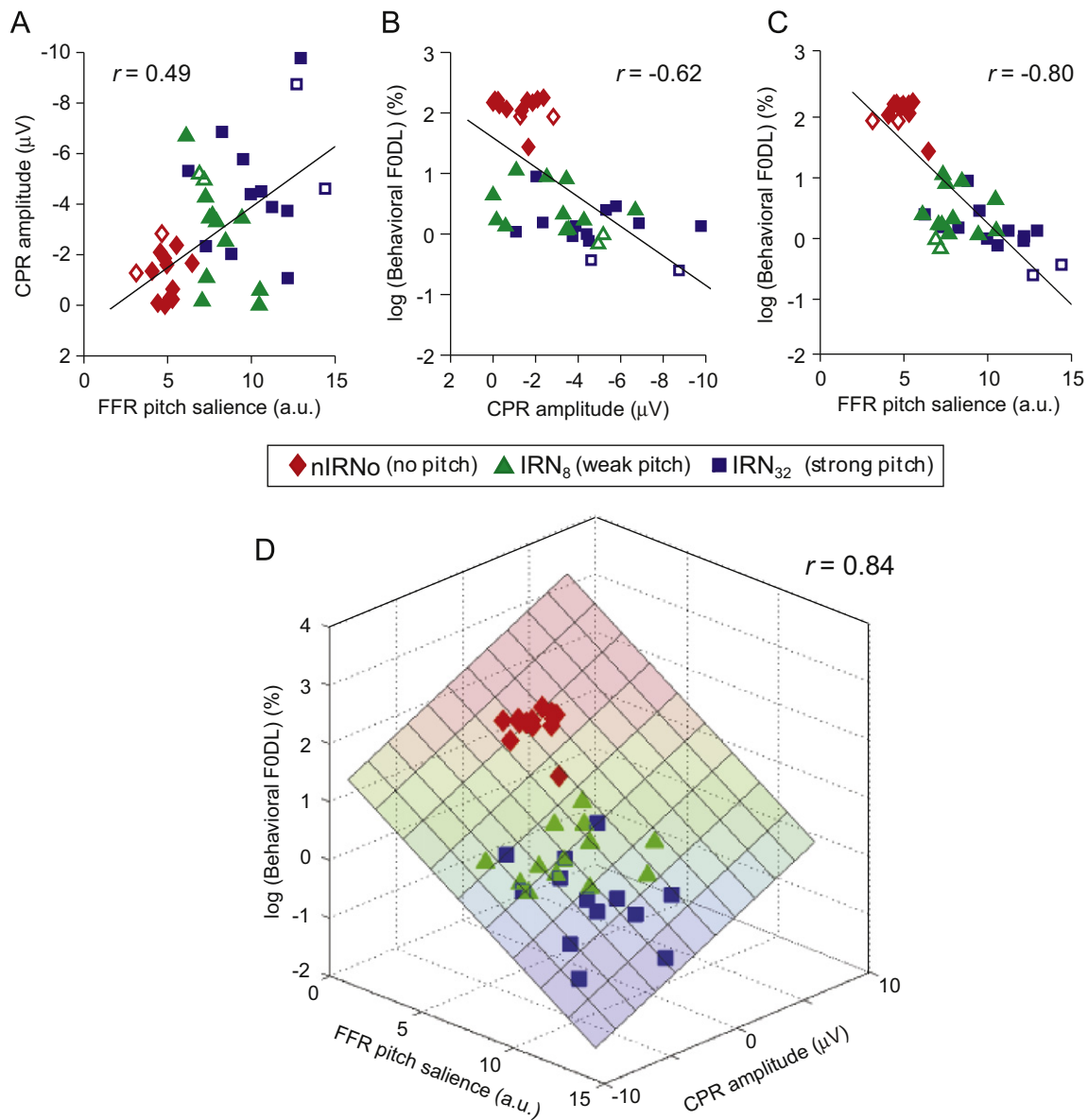
The major findings of this study show that neural pitch strength (as reflected in the brainstem FFR and the cortical CPR) and behavioral pitch salience (as reflected by the  $F_0$  DL estimates) improve systematically with increasing temporal regularity of the IRN stimulus. The strong correlation between neural representations relevant to pitch salience at the brainstem and early cortical levels of processing suggests that sensory processing at the brainstem level may be driving early preattentive sensory processing relevant to pitch at the cortical level. In addition, the strong correlation between the behavioral  $F_0$  DL measure and both neural measures of pitch suggests that early sensory level processing of pitch-relevant information in the brainstem and auditory cortex plays an important role in shaping pitch perception.

##### 4.1. Pitch-relevant information preserved in the FFR and CPR is sensitive to changes in pitch salience

In agreement with our previous report on FFR responses to IRN stimuli varying in temporal regularity (Krishnan et al., 2010), our findings herein similarly show greater FFR pitch salience magnitude for a stimulus with stronger pitch salience. We conclude that the improving temporal regularity of the IRN stimulus with increasing iteration steps increases the degree of neural phase-locking to the pitch-relevant periodicity in the stimulus. This view is compatible with previous perceptual (Patterson et al., 1996; Yost, 1996b) and physiologic (Bilsen, ten Kate, Buunen, & Raatgever, 1975; Sayles & Winter, 2007; Shofner, 1991; Shofner, 1999; Winter, Wiegrebe, & Patterson, 2001) data suggesting that pitch of static and dynamic IRN stimuli is based on autocorrelation-like temporal processing. Specifically, coincidence detection neurons in the IC perform a correlation analysis on the delayed and undelayed temporal information from the cochlear nucleus to extract pitch-relevant periodicities that are spatially mapped onto a periodicity pitch axis. By this analysis, the magnitude of the autocorrelation peak increases with iteration step, and provides a strong measure of temporal regularity, a stimulus attribute closely tied to pitch salience.

Similarly, we observe greater CPR magnitude (and shorter response latency) for IRN stimuli with stronger pitch. This finding is consistent with previous reports that have shown that the latency and amplitude of POR varies systematically with the pitch salience of an IRN stimulus (Krumbholz et al., 2003; Seither-Preisler et al., 2006; Soeta, Nakagawa, & Tonoike, 2005). It is suggested that the neural activity underlying the generation of the pitch onset response is involved in extracting an initial estimate of the pitch salience of the sound. In addition, functional imaging studies in humans show that activity of the primary auditory cortex increases as a function of the number of iteration steps (Griffiths et al., 1998; Griffiths, Uppenkamp, Johnsrude, Josephs, & Patterson, 2001). Intracranial electrode recordings reveal that discharge rates increase in auditory cortical neurons as a function of iteration steps in both primates (Bendor & Wang, 2005) and humans (Schonwiesner & Zatorre, 2008). That is, unlike the brainstem where periodicity and pitch are often represented by regular temporal patterns of action potentials that are phase-locked to the sound waveform, the most commonly observed code for periodicity and pitch within cortical neurons is a modulation of spike rates as a function of  $F_0$ . It is possible that the wider temporal integration window at the cortical level may render the auditory cortical neurons too sluggish to provide phase-locked representations of periodicity within the pitch range (Walker, Bizley, King, & Schnupp, 2011). However, it is not yet clear how cortical neurons transform the autocorrelation-like temporal analysis in the brainstem to a spike rate code to extract pitch-relevant information. Nevertheless, these findings collectively suggest that the increase in pitch salience with increasing temporal regularity of the IRN stimulus is correlated with an increase in pitch-relevant neural activity in both brainstem and cortical auditory neurons.

Our results show a strong correlation in neural activity relevant to pitch salience between the brainstem and the auditory cortex recorded concurrently. This finding suggests that brainstem and cortical representations of pitch are shaped in a coordinated manner through corticofugal modulation of subcortical afferent circuitry (Nahum, Nelken, & Ahissar, 2008). The notion that neural processes at the brainstem and cortical levels may be well coordinated is supported by neuroanatomical evidence for ascending and descending pathways (Huffman & Henson, 1990; Kelly & Wong, 1981; Saldana et al., 1996) and physiologic evidence for cortical modulation of brainstem



**Fig. 9.** Comparison of neural and behavioral measures of pitch salience. Pair-wise regressions between neural and perceptual measures: (A) brainstem FFR vs. cortical CPR; (B) CPR vs. behavioral  $F_0$  DLs; (C) FFR vs.  $F_0$  DLs. In panels A–C, observe that stimuli with no pitch (IRNO; red diamonds) or weak pitch salience (IRN8; green triangles) cluster away from stimuli with high pitch salience (IRN32; blue squares). Also, note the open symbols for data points from the two trained musicians. (D) Simultaneous representation of pitch in brainstem (FFR) and early cortical processing (CPR) is associated with poorer perceptual performance (i.e., larger DLs; upper portion of the plane), whereas stronger neural representations at both levels are associated with improved performance (i.e., lower DLs; lower portion of the plane). These correlations suggest that a listener's pitch discrimination ability is predicted, at least in part, by how well pitch cues are represented in the brainstem and the auditory cortex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

representations (Suga et al., 2003; Yan & Suga, 1998; Zhou & Jen, 2000). In the case of humans, the corticofugal system likely shapes the enhancement of brainstem pitch representation resulting from short-term auditory training (Russo et al., 2005; Song et al., 2008); long-term linguistic experience (Krishnan & Gandour, 2009; Krishnan et al., 2012, 2005); and musical training (Bidelman et al., 2011a; Bidelman & Krishnan, 2009; Musacchia et al., 2007; Wong et al., 2007). The reverse hierarchy theory (RHT) provides a representational hierarchy to describe the interaction between sensory input and top-down processes in primary sensory areas (Ahissar & Hochstein, 2004; Nahum et al., 2008; Shamma, 2008). Its basic claim is that neural circuitry mediating a certain percept can be modified starting at the highest representational level and progressing to lower-levels in search of more refined high resolution information to optimize

perception. RHT has been invoked as a plausible explanation for top-down influences on subcortical sensory processing (Banai, Abrams, & Kraus, 2007). Consistent with this theory, early sensory level representation of spectro-temporal features important for pitch in the brainstem would be more precise compared to the early sensory level representation of pitch-relevant information in the auditory cortex which could be more labile and spatiotemporally broader (Warren & Griffiths, 2003; Zatorre & Belin, 2001; Winer, Miller, Lee, & Schreiner, 2005; Chechik, Anderson, Bar-Yosef, Young & Tishby, 2006).

#### 4.2. Is the CPR a pitch-specific response?

Using IRN and IRNO stimuli, we observe a robust CPR to IRN pitch-eliciting stimuli (IRN8, IRN32) but not to IRNO. These data



confirm that the CPR is indeed specific to pitch rather than simply a neural response to IRN elicited by slow, spectro-temporal modulations that are unrelated to pitch. As such, they are inconsistent with data from functional imaging studies of pitch showing that both IRN and IRN0 stimuli yield similar response patterns within lateral Heschl's Gyrus and across auditory cortex (Barker et al., 2012; Ernst, Verhey, & Uppenkamp, 2008; Hall & Plack, 2009). Nevertheless, what is clear is that the spectral flux of IRN, by itself, is insufficient to account for the fact that the CPR is a pitch-specific response.

Several reasons might account for the disparity in findings between fMRI and EEG/MEG experiments. Evoked potentials arguably provide a more direct assay of neuronal excitation or inhibition than the fMRI blood oxygen level dependent response. This is because the latter depends on hemodynamic changes in blood flow in neural tissue that provide only an indirect measure of underlying neural activity. As other neuroimaging methodologies (e.g., PET, MEG) have demonstrated a pitch sensitive response to IRN, it is also possible that there is a specific confound related to fMRI. As pointed out by Ernst et al. (2008), both the limited spatiotemporal resolution and imaging noise of fMRI may prevent the possibility of observing systematic changes in response sensitivity to detailed features of an auditory signal. Indeed, fMRI may only reveal auditory regions which are maximally responsive to either the presence or absence of pitch instead of its underlying perceptual dimensions (Barker, Plack, & Hall, 2011, p. 115). The failure of fMRI to identify representations for pitch salience in either subcortical (Hall & Plack, 2009; Penagos et al., 2004) or cortical (Barker et al., 2011) structures may in part be due to the fact that representations of pitch or pitch salience exist on a much finer temporal scale (milliseconds) than that (seconds) resolvable by hemodynamic imaging methods. In contrast, the identification of neural correlates of pitch salience in the present study corroborates previous work demonstrating that electrophysiological recording techniques (EEG/MEG) provide robust methods to examine the neural correlates of complex pitch and its various dimensions (Chait et al., 2006; Krishnan et al., 2010; Krishnan et al., 2010; Krumbholz et al., 2003; Schonwiesner & Zatorre, 2008).

#### 4.3. Comparison of neural and behavioral representations of pitch relevant information

Our findings reveal a strong correlation between the neural (FFR, CPR) and behavioral ( $F_0$  DL) measures of pitch salience. These data are in agreement with previous studies of the relationship between neural and behavioral measures of pitch: brainstem (Bidelman & Krishnan, 2011; Krishnan et al., 2010; Krumbholz et al., 2003); cerebral cortex (Krumbholz et al., 2003; Soeta et al., 2005). We therefore infer that the growth in pitch salience with increasing temporal regularity may be explained by improved neural representation of pitch-relevant periodicities in the brainstem response, and by the shorter latency and greater amplitude of the cortical pitch response. The latter finding suggests improved temporal synchronization of pitch-relevant neural activity among cortical neurons generating the CPR. The strong correlation observed herein between the neural and behavioral measures supports the view that neural representation of pitch relevant information at early sensory stages (brainstem and auditory cortex) of processing plays an important role in shaping pitch perception.

Our findings further show that the correlation between the brainstem pitch representation and the behavioral pitch measure was stronger than that observed between the cortical pitch representation and the behavioral measure. Yet in this experiment we are unable to determine whether these differences

reflect greater variability associated with the more labile cortical response or a fundamental difference in the fidelity of the representations at the two levels. Nonetheless, the comparatively stronger correlation between the brainstem pitch representation and the behavioral measure of pitch is consistent with RHT. We therefore speculate that cortical processes contributing to pitch perception may put more weight on the more precise brainstem representation of pitch relative to the less precise and more labile early sensory level cortical representation in order to optimize perception.

#### 4.4. Conclusions

A novel experimental paradigm has been developed to acquire brainstem and cortical pitch-relevant responses concurrently that enables us to evaluate neural representation of pitch salience in the human brainstem and auditory cortex, and its relationship to a perceptual measure of pitch salience. Perceptual pitch salience of IRN stimuli, as reflected by the degree of stimulus periodicity, can be predicted by brainstem and/or cortical neural representations of pitch-relevant information. These findings support the view that neural representations of pitch information germane to the development of pitch percept is already extracted by early preattentive sensory level processing in both the brainstem and auditory cortex. Future work is necessary to fully characterize properties of the CPR elicited by experience-dependent, ecologically-relevant pitch patterns in the language (e.g., dynamic, curvilinear pitch) and music domains (e.g., consonance/dissonance). This paradigm may also provide a new window to evaluate the interplay online between feedforward and feedback components in the processing of pitch-relevant information at the level of the brainstem and the auditory cortex (cf. Foxe & Schroeder, 2005).

#### Acknowledgments

Research supported by NIH R01 DC008549 (A.K.), NIDCD T32 DC00030 (C.J.S.), and Purdue University Bilsland Dissertation Fellowship (G.B.). We thank Dr. Chris Plack (University of Manchester) for sharing the MATLAB script used to generate the IRN0 stimuli. Reprint requests should be addressed to Ananthanarayan Krishnan, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907-2038, USA, or via email: rkish@purdue.edu.

#### References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8(10), 457–464.
- ASA. (1960). Acoustical terminology SI, pp. 1–1960. American Standards Association.
- Banai, K., Abrams, D., & Kraus, N. (2007). Sensory-based learning disability: insights from brainstem processing of speech sounds. *International Journal of Audiology*, 46(9), 524–532.
- Barker, D., Plack, C. J., & Hall, D. A. (2011). Human auditory cortical responses to pitch and to pitch strength. *Neuroreport*, 22, 111–115.
- Barker, D., Plack, C. J., & Hall, D. A. (2012). Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. *Cerebral Cortex*, 22, 745–753.
- Baumann, S., Griffiths, T. D., Sun, L., Petkov, C. I., Thiele, A., & Rees, A. (2011). Orthogonal representation of sound dimensions in the primate midbrain. *Nature Neuroscience*, 14, 423–425.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436, 1161–1165.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, 23, 425–434.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, 77, 1–10.

- Bidelman, G. M., & Heinz, M. G. (2011). Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *Journal of the Acoustical Society of America*, 130, 1488–1502.
- Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, 29, 13165–13171.
- Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, 22, 212–216.
- Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European Journal of Neuroscience*, 33, 530–538.
- Bilsen, F. A., ten Kate, J. H., Buunen, T. J. F., & Raatgever, J. (1975). Responses of single units in the cochlear nucleus of the cat to cosine noise. *Journal of the Acoustical Society of America*, 58, 858–866.
- Cariani, P. A., & Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology*, 76, 1698–1716.
- Cedolin, L., & Delgutte, B. (2005). Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. *Journal of Neurophysiology*, 94, 347–362.
- Chait, M., Poeppel, D., & Simon, J. Z. (2006). Neural response correlates of detection of monaurally and binaurally created pitches in humans. *Cerebral Cortex* (New York, N.Y.: 1991), 16, 835–848.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology*, 47, 236–246.
- Chechik, G., Anderson, M. J., Bar-Yosef, O., Young, E. D., Tishby, N., et al. (2006). Reduction of information redundancy in the ascending auditory pathway. *Neuron*, 51, 359–368.
- Demany, L., & Semal, C. (2002). Learning to perceive pitch differences. *Journal of the Acoustical Society of America*, 111, 1377–1388.
- Ernst, S. M., Verhey, J. L., & Uppenkamp, S. (2008). Spatial dissociation of changes of level and signal-to-noise ratio in auditory cortex for tones in noise. *Neuroimage*, 43, 321–328.
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16, 419–423.
- Galbraith, G., Threadgill, M., Hemsley, J., Salour, K., Songdej, N., Ton, J., et al. (2000). Putative measure of peripheral and brainstem frequency-following in humans. *Neuroscience Letters*, 292, 123–127.
- Goldstein, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 54, 1496–1516.
- Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1, 422–427.
- Griffiths, T. D., Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, 20, 1128–1132.
- Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, 4, 633–637.
- Gutschalk, A., Patterson, R. D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage*, 15, 207–216.
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *Neuroimage*, 22, 755–766.
- Hall, D. A., & Plack, C. J. (2009). Pitch processing sites in the human auditory brain. *Cerebral Cortex*, 19, 576–585.
- Huffman, R. F., & Henson, O. W., Jr. (1990). The descending auditory pathway and acousticomotor systems: connections with the inferior colliculus. *Brain Research. Brain Research Reviews*, 15, 295–323.
- Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, 123, 155–163.
- Kelly, J. P., & Wong, D. (1981). Laminar connections of the cat's auditory cortex. *Brain Research*, 212, 1–15.
- Kraus, N., & Banai, K. (2007). Auditory-processing malleability: focus on language and music. *Current Directions in Psychological Science*, 16, 105–110.
- Krishnan, A. (2007). Human frequency following response. In: R. F. Burkard, M. Don, & J. J. Eggermont (Eds.), *Auditory evoked potentials: basic principles and clinical application* (pp. 313–335). Baltimore: Lippincott Williams & Wilkins.
- Krishnan, A., Bidelman, G. M., & Gandour, J. T. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, 268, 60–66.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, 110, 135–148.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2012). Experience-dependent plasticity in pitch encoding: from brainstem to auditory cortex. *Neuroreport*, 23, 498–502.
- Krishnan, A., Gandour, J. T., Smalt, C. J., & Bidelman, G. M. (2010). Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain and Language*, 114, 193–198.
- Krishnan, A., & Plack, C. J. (2009). Auditory brainstem correlates of basilar membrane nonlinearity in humans. *Audiology and Neuro-Otology*, 14, 88–97.
- Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, 21, 1092–1105.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research. Cognitive Brain Research*, 25, 161–168.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13, 765–772.
- Langner, G. (1983). Evidence for neuronal periodicity detection in the auditory system of the Guinea fowl: implications for pitch analysis in the time domain. *Experimental Brain Research*, 52, 333–355.
- Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience*, 29, 5832–5840.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467–477.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219, 36–47.
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 15894–15898.
- Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, 241, 34–42.
- Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: the case of speech in noise. *PLoS Biology*, 6(5), e126.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Pantev, C., Hoke, M., Lutkenhoner, B., & Lehnertz, K. (1989). Tonotopic organization of the auditory cortex: pitch versus frequency representation. *Science*, 246, 486–488.
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, 30, 653–661.
- Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, 11, 369–372.
- Patterson, R. D., Handel, S., Yost, W. A., & Datta, A. J. (1996). The relative strength of the tone and noise components in iterated ripple noise. *Journal of the Acoustical Society of America*, 100, 3286–3294.
- Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, 24, 6810–6815.
- Plack, C. J. (2005). *Pitch: neural coding and pitch perception*. New York: Springer.
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, 156, 95–103.
- Saldana, E., Feliciano, M., & Mugnaini, E. (1996). Distribution of descending projections from primary auditory neocortex to inferior colliculus mimics the topography of intracollicular projections. *Journal of Comparative Neurology*, 371, 15–40.
- Sayles, M., & Winter, I. M. (2007). The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Research*, 1171, 52–66.
- Schonwiesner, M., & Zatorre, R. J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, 187, 97–105.
- Seither-Preisler, A., Patterson, R. D., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). From noise to pitch: transient and sustained responses of the auditory evoked field. *Hearing Research*, 218, 50–63.
- Shamma, S. (2008). On the emergence and awareness of auditory objects. *PLoS Biology*, 6(6), e155.
- Shofner, W. P. (1991). Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. *Journal of the Acoustical Society of America*, 90, 2450–2466.
- Shofner, W. P. (1999). Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural autocorrelograms. *Journal of Neurophysiology*, 81, 2662–2674.
- Soeta, Y., Nakagawa, S., & Tonoike, M. (2005). Auditory evoked magnetic fields in relation to iterated rippled noise. *Hearing Research*, 205, 256–261.
- Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, 20, 1892–1902.
- Suga, N., Ma, X., Gao, E., Sakai, M., & Chowdhury, S. A. (2003). Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. *Speech Communication*, 41, 189–200.
- Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for the extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71, 679–687.
- Walker, K. M. M., Bizley, J. K., King, A. J., & Schnupp, J. W. H. (2011). Cortical encoding of pitch: recent results and open questions. *Hearing Research*, 271, 74–87.
- Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience*, 23, 5799–5804.
- Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends in Neuroscience*, 28, 255–263.

- Winter, I. M., Wiegand, L., & Patterson, R. D. (2001). The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *Journal of Physiology*, *537*, 553–566.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.
- Yan, W., & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, *1*, 54–58.
- Yost, W. A. (1996a). Pitch of iterated rippled noise. *Journal of the Acoustical Society of America*, *100*, 511–518.
- Yost, W. A. (1996b). Pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America*, *100*, 3329–3335.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, *84*, 566–572.
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*, 946–953.
- Zhou, X., & Jen, P. H. (2000). Brief and short-term corticofugal modulation of subcortical auditory responses in the big brown bat, *Eptesicus fuscus*. *Journal of Neurophysiology*, *84*, 3083–3087.