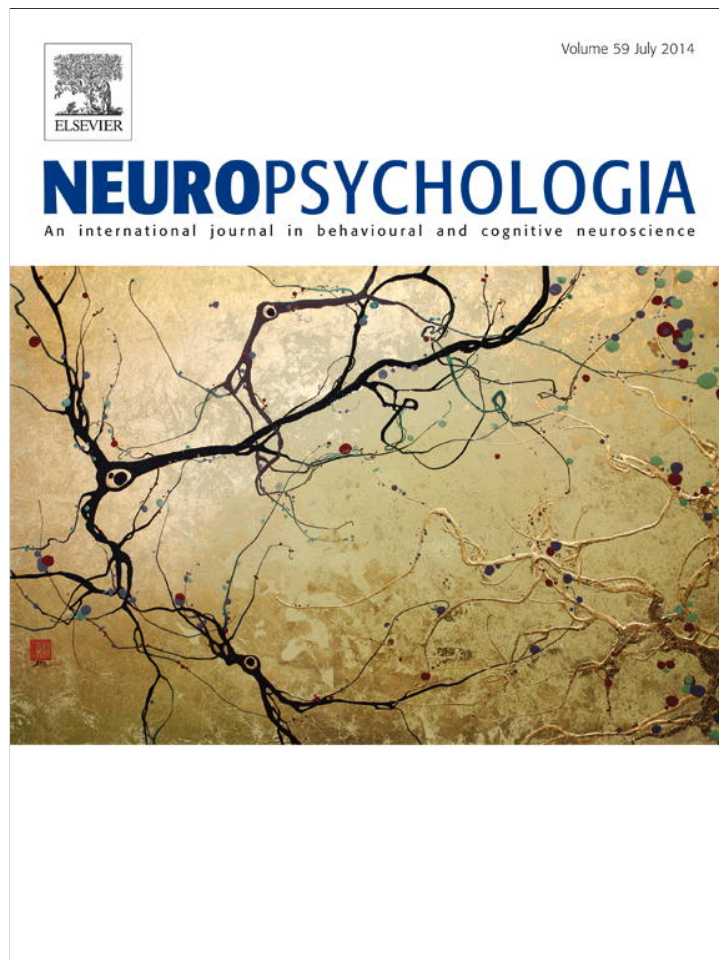


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

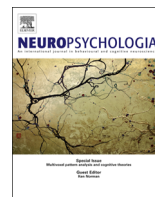
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at ScienceDirect

Neuropsychologia

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

Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours



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

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

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

ARTICLE INFO

Article history:

Received 9 November 2013

Received in revised form

12 March 2014

Accepted 11 April 2014

Available online 18 April 2014

Keywords:

Auditory

Pitch encoding

Iterated rippled noise

Cortical pitch response

Pitch onset response

Hemispheric laterality

ABSTRACT

Voice pitch is an important information-bearing component of language that is subject to experience dependent plasticity at both early cortical and subcortical stages of processing. We have already demonstrated that pitch onset component (Na) of the cortical pitch response (CPR) is sensitive to *flat* pitch and its salience. In regards to *dynamic* pitch, we do not yet know whether the multiple pitch-related transient components of the CPR reflect specific temporal attributes of such stimuli. Here we examine the sensitivity of the multiple transient components of CPR to changes in pitch acceleration associated with the Mandarin high rising lexical tone. CPR responses from Chinese listeners were elicited by three citation forms varying in pitch acceleration and duration. Results showed that the pitch onset component (Na) was invariant to changes in acceleration. In contrast, Na–Pb and Pb–Nb showed a systematic increase in the interpeak latency and decrease in amplitude with increase in pitch acceleration that followed the time course of pitch change across the three stimuli. A strong correlation with pitch acceleration was observed for these two components only – a putative index of pitch-relevant neural activity associated with the more rapidly-changing portions of the pitch contour. Pc–Nc marks unambiguously the stimulus offset. We therefore propose that in the early stages of cortical sensory processing, a series of neural markers flag different temporal attributes of a dynamic pitch contour: onset of temporal regularity (Na); changes in temporal regularity between onset and offset (Na–Pb, Pb–Nb); and offset of temporal regularity (Pc–Nc). At the temporal electrode sites, the stimulus with the most gradual change in pitch acceleration evoked a rightward asymmetry. Yet within the left hemisphere, stimuli with more gradual change were indistinguishable. These findings highlight the *emergence* of early hemispheric preferences and their functional roles as related to sensory and cognitive properties of the stimulus.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Pitch is an important information-bearing perceptual attribute that plays an important role in the perception of language and music. There is considerable interest therefore in how this pitch-relevant information is extracted from speech and nonspeech sounds at both subcortical (Cariani & Delgutte, 1996a, 1996b; Cedolin & Delgutte, 2005; Meddis & O'Mard, 1997) and cortical levels (Walker, Bizley, King, & Schnupp, 2011). There is also growing interest in understanding the neural mechanisms that mediate

experience-dependent shaping of pitch processing. Linguistic and musical pitch provide us with a window to evaluate how neural representation of pitch-relevant attributes emerge from early sensory levels of processing and interact with higher levels of cognitive processing in the human brain, and how language and music experience shapes these representations. Recent empirical data show that neural representation of pitch is shaped by one's experience with language and music at the level of the auditory brainstem as well as the cerebral cortex (Besson, Chobert & Marie, 2011; Gandour & Krishnan, 2014; Koelsch, 2012; Kraus & Banai, 2007; Krishnan, Gandour, & Bidelman, 2012; Meyer, 2008; Munte, Altenmuller, & Jancke, 2002; Patel & Iversen, 2007; Tervaniemi et al., 2009; Zatorre, Belin, & Penhune, 2002; Zatorre & Gandour, 2008). But we have yet to achieve a precise characterization of neural representation of pitch-relevant information associated with specific attributes of *dynamic* pitch contours that occur in natural speech.

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

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

At the cortical level, magnetoencephalography (MEG) has been used previously to study the sensitivity to periodicity, an essential requisite of pitch, by investigating the N100m component. However, a large proportion of the N100m is simply a response to the onset of sound energy, and not exclusively to pitch (Alku, Sivonen, Palomaki, & Tiitinen, 2001; Gutschalk, Patterson, Scherg, Uppenkamp, & Rupp, 2004; Lutkenhoner, Seither-Preisler, & Seither, 2006; Soeta & Nakagawa, 2008; Yrttiaho, Tiitinen, May, Leino, & Alku, 2008). In order to disentangle the pitch-specific response from the onset response, a novel stimulus paradigm was constructed with two segments – an initial segment of noise with no pitch to evoke the onset components only, followed by a pitch-eliciting segment of iterated rippled noise (IRN) matched in intensity and overall spectral profile (Krumbholz, Patterson, Seither-Preisler, Lammertmann, & Lutkenhoner, 2003). Interestingly, a transient pitch onset response (POR) was evoked from this noise-to-pitch transition only. The reverse stimulus transition from pitch to noise failed to produce a POR. It has been proposed that the human POR, as measured by MEG, reflects synchronized cortical neural activity specific to pitch (Chait, Poeppel, & Simon, 2006; Krumbholz et al., 2003; Ritter, Gunter Dosch, Specht, & Rupp, 2005; Seither-Preisler, Patterson, Krumbholz, Seither, & Lutkenhoner, 2006). POR latency and magnitude, for example, has been shown to depend on pitch salience. A more robust POR with shorter latency is observed for stimuli with stronger pitch salience as compared to those with weaker pitch salience. Source analyses (Gutschalk, Patterson, Rupp, Uppenkamp, & Scherg, 2002; Gutschalk et al., 2004; Krumbholz et al., 2003), corroborated by human depth electrode recordings (Griffiths et al., 2010; Schonwiesner & Zatorre, 2008), indicate that the POR is localized to the anterolateral portion of Heschl's gyrus, the putative site of pitch processing (Bendor & Wang, 2005; Griffiths, Buchel, Frackowiak, & Patterson, 1998; Johnsrude, Penhune, & Zatorre, 2000; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Penagos, Melcher, & Oxenham, 2004; Zatorre, 1988).

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

In MEG/EEG studies of pitch processing, the primary focus was to characterize a single, transient pitch onset response to stimuli exhibiting a steady-state pitch. Whether or not their findings can be extrapolated to neural mechanisms that are associated with dynamic pitch contours is an empirical question. In this study, we utilize linguistically-relevant pitch stimuli with dynamic, curvilinear patterns. They represent *within-category* variants of the high rising lexical tone of Mandarin Chinese.

Linguistic and musical pitch provide an analytic window to evaluate how neural representations of important pitch attributes of a sound undergo transformation from early sensory to later cognitive stages of processing in the human brain, and how pitch-relevant experience shapes these representations. In the music domain, MEG/EEG studies have shown enhanced auditory cortical representations (Pantev et al., 1998; Shahin, Bosnyak, Trainor, & Roberts, 2003) and superior detection of weak pitch incongruities

(Magne, Schon, & Besson, 2006; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011) in musicians compared with nonmusician controls. Though providing convincing evidence in support of experience-dependent neural plasticity, none of their evoked response components were pitch-dependent. In the language domain, tonal languages are especially advantageous for studying pitch processing because of its functional load at the level of the syllable. Almost all previous EEG studies of lexical tone have investigated mismatch negativity (MMN) responses by using dynamic pitch contours with a passive oddball paradigm (cf. Gu, Zhang, Hu, Zhao, & Zhang, 2013; steady-state pitch; Zheng, Minett, Peng, & Wang, 2012; active oddball, P300). As measured by MMN amplitude, early, preattentive processing of lexical tone, relative to consonants, was lateralized to the right hemisphere (RH) (Luo et al., 2006). This RH preference was also observed in an MMN study of categorical perception of lexical tone (Xi, Zhang, Shu, Zhang, & Li, 2010). In their nonspeech condition that differed only in spectral components from the speech condition, no difference was observed between within- and between-category deviants, relative to standard stimuli, in the RH. Of special relevance to this study, *within-category* deviants elicited larger MMN in the RH regardless of speech/nonspeech condition. Lexical tone is of further interest because its primary auditory correlate is based on variations in pitch, a multidimensional perceptual attribute that relies on several acoustic features. MMN studies of tone languages have revealed that pitch contour and pitch height are two important features used in early, preattentive lexical tone processing (Chandrasekaran, Gandour, & Krishnan, 2007; Chandrasekaran, Krishnan, & Gandour, 2007; Tsang, Jia, Huang, & Chen, 2011). These findings notwithstanding, any definitive interpretations about cortical pitch processing must be tempered by the fact that the MMN is comprised of both auditory and cognitive mechanisms of frequency change detection in auditory cortex (Maess, Jacobsen, Schroger, & Friederici, 2007). The MMN itself is not a pitch-specific response.

Our overall objective therefore is to fill this knowledge gap by examining cortical, pitch-specific responses that are elicited by linguistically-relevant, dynamic pitch contours exemplary of those that occur in natural speech. To our knowledge there are no published reports that tie specific transient components of the cortical pitch response to selected portions of dynamic pitch contours. The specific aim of this paper is to first identify the multiple transient components of the CPR and label them in relation to specific aspects of our dynamic, curvilinear pitch stimuli (e.g., pitch onset, pitch offset, pitch acceleration). Our hypothesis is that the cortical representation of pitch, as reflected by the CPR, will be differentially sensitive to different attributes or features of dynamic pitch. To eliminate any potential interference of higher-order phonological categories, all three of our dynamic pitch stimuli fall within the limits of citation forms of Mandarin Tone 2. To enable us to focus primarily on the effects of changes in rate of acceleration during the rising portion of Tone 2, pitch height is constant; i.e., pitch onset and offset and pitch range are identical across stimuli. With respect to duration, the three stimuli represent short-, medium-, and long-duration variants of isolated productions of Tone 2.

2. Materials and methods

2.1. Participants

Ten native speakers of Mandarin Chinese (5 males, 5 females) were recruited from the Purdue University student body to participate in the experiment. All exhibited normal hearing sensitivity at audiometric frequencies between 500 and 4000 Hz and reported no previous history of neurological or psychiatric illnesses. They were closely matched in age (24.50 ± 3.53 years), years of formal education (16.90 ± 2.88 years), and were strongly right handed (laterality index $93.10 \pm 11.22\%$)

as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). All participants were born and raised in mainland China. None had received formal instruction in English before the age of nine (12 ± 1.24 years). As determined by a music history questionnaire (Wong & Perrachione, 2007), all participants, except for one, had less than three years of musical training (1.55 ± 3.00 years) on any combination of instruments. None had any training within the past five years. Each participant was

paid and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

2.2. Stimuli

Using an iterated rippled noise (IRN) algorithm, dynamic pitch contours approximating three isolated, citation variants of Mandarin Tone 2 were constructed: short (T2_150), intermediate (T2_200), and long (T2_250). Their durations were, in order, 150, 200, and 250 ms (Appendix A: audio, stimuli). Though infrequent, a short variant (T2_150) has been reported to occur in isolated productions of Tone 2 (Kratochvil, 1985). These stimuli differed in F_0 rate of acceleration as well as duration (Fig. 1). These durations easily fall outside the range of temporal integration effects (≈ 80 ms) on pitch and its salience for stimuli with resolved harmonics (Plack, Carlyon, & Viemeister, 1995; Plack, Turgeon, Lancaster, Carlyon, & Gockel, 2011; Plack & White, 2000; White & Plack, 1998, 2003). It is therefore unlikely that temporal integration effects pose a potential confound for our evaluation of pitch acceleration-related effects. Rates of acceleration are displayed at 80 ms and from minimum to maximum (in the acceleration domain) per stimulus in Table 1. The maximum speed of pitch change within a speaker's ability to produce a rapid shift in rising pitch over a 4 st interval is 61.3 st/s (Xu & Sun, 2002, p. 1407, Table VII). The average velocity rates (in st/s), calculated from the turning point to F_0 offset, for T2_250 (25.6), T2_200 (32.1), and T2_150 (42.7) fall within the physiological limits of speed of rising pitch changes. As reflected by FFR responses in the brainstem (Krishnan, Gandour, Smalt, & Bidelman, 2010, p. 96, Figs. 2 and 3), a scaled variant of Tone 2 with a velocity rate of 51.94 st/s falls within the bounds of the normal voice range. F_0 onset (100.88 Hz) and offset (131.72 Hz) were constant. ΔF_0 from turning point to offset was fixed across stimuli at 30.84 Hz (4.6 st; .38 octaves). This ΔF_0 value is comparable to that of an exemplary Tone 2 citation form (Krishnan et al., 2010) and is an effective cue for the perception of isolated Tone 2 (Moore & Jongman, 1997). The turning point was located at about $\approx 26\%$ of the duration of the F_0 contour (40 ms, T2_150; 53 ms, T2_200; 66 ms, T2_250). The timing of these turning points relative to F_0 onset are perceptually relevant in the identification of Tone 2 (cf. Moore & Jongman, 1997, p. 1870, Fig. 4). Based on these behavioral and neural data, we judged these stimuli to be ecologically valid representations of Tone 2 and moreover likely to elicit differential sensitivity to varying degrees of acceleration rates at the cortical level.

These three IRN stimuli with time-varying F_0 contours were generated by applying a time-varying, delay-and-add algorithm (Appendix B: text, fourth-order polynomial equations) (Denham, 2005; Krishnan, Swaminathan, & Gandour, 2009;

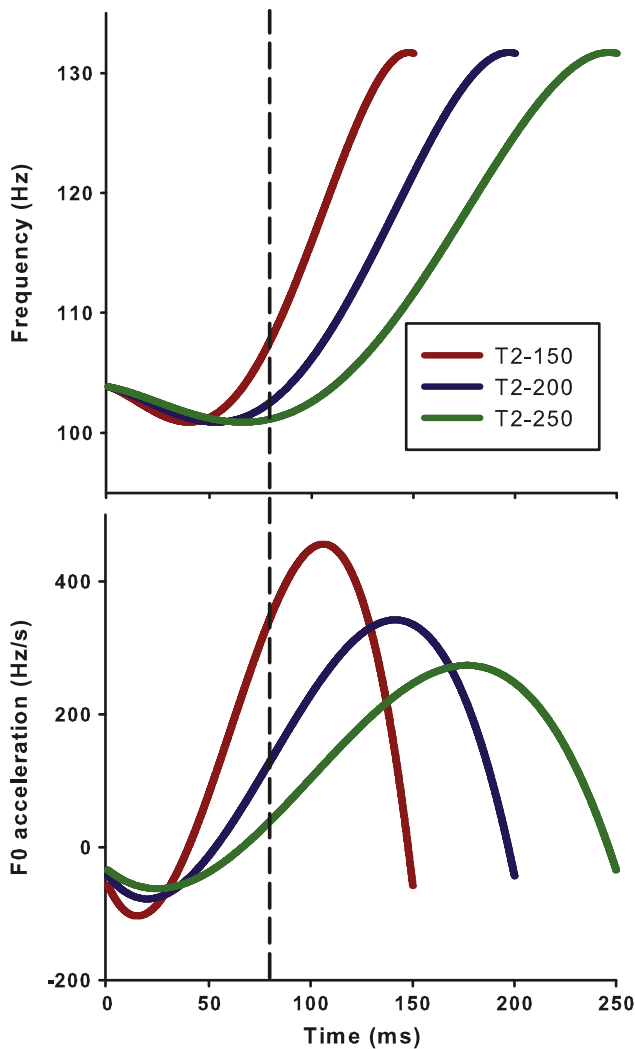


Fig. 1. IRN stimuli used to evoke cortical responses to linguistic patterns that are differentiated by varying degrees of rising acceleration and duration. Voice fundamental frequency (F_0) contours (top panel) and corresponding acceleration trajectories (bottom panel) of all three stimuli are modeled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

Table 1
 F_0 acceleration rates per stimulus.

Stimulus	Instantaneous ^a		Minimum to maximum ^b	
	Hz/s	st/s	Hz/s	st/s
T2_150	346	56	559	85
T2_200	130	22	419	63
T2_250	40	7	336	51

^a Located at 80 ms within pitch segment.
^b Average.

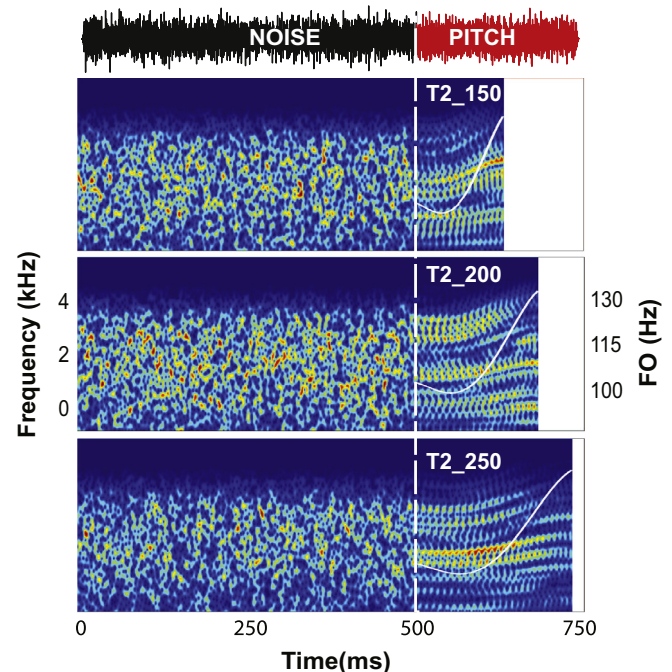


Fig. 2. Waveform (T2_250) and spectrograms of each of the three stimulus conditions (T2_150; T2_200; T2_250) illustrate the experimental paradigm used to acquire cortical responses. The vertical dashed line at 500 ms demarcates the transition from the initial noise segment to the final pitch segment. FFRs and CPRs were extracted from evoked responses beginning with the onset of the pitch. F_0 contours (white) are superimposed on their respective pitch segments. Within the pitch segment, the waveform (top) shows robust periodicity at a high IRN iteration step ($n=32$); the spectrograms show clear resolution of dynamic, rising spectral bands corresponding to the harmonics of the fundamental frequency.

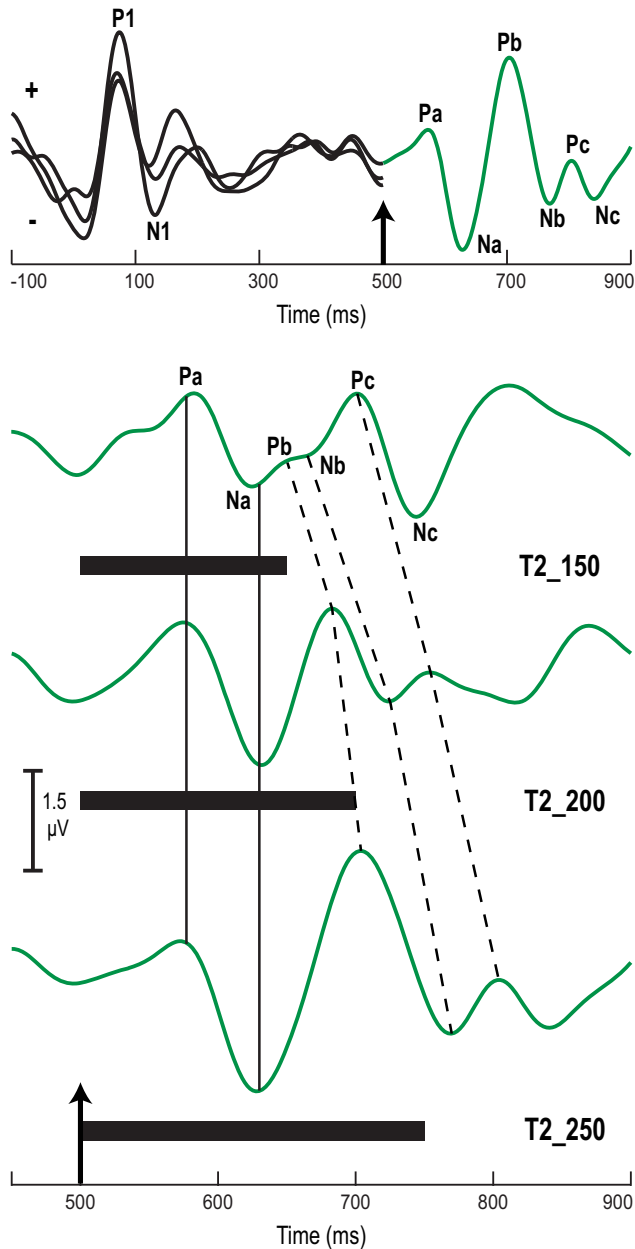


Fig. 3. Grand averaged cortical evoked response components at the Fz electrode site per stimulus condition. The P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (green) are displayed in the top panel. The up arrow at 500 ms marks the boundary between the noise precursor segment and the pitch-eliciting stimulus. Na, Pb, and Nb are the most robust response components. CPR waveforms (green) elicited by the three stimuli (T2_150, T2_200, T2_250) are shown in the bottom panel. The up arrow at 500 ms marks the beginning of the pitch-eliciting stimulus. Solid black horizontal bars indicate the duration of each stimulus. Whereas Pa and Na do not change across stimuli (solid vertical lines), Pb, Nb, Pc, and Nc all show a systematic increase in peak latency (dashed vertical lines). Response amplitude for Na, Pb, and Nb increases from T2_150 to T2_250 in conjunction with decreasing pitch acceleration and increasing duration across stimuli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Sayles & Winter, 2007; Swaminathan, Krishnan, Gandour, & Xu, 2008). A high iteration step ($n=32$) was chosen because pitch salience does not increase by any noticeable amount beyond this number of iteration steps. The gain was set to 1. By using IRN, we preserve dynamic variations in pitch of auditory stimuli that lack a waveform periodicity, formant structure, temporal envelope, and recognizable timbre characteristic of speech.

Each stimulus condition consisted of two segments (crossfaded with 5 ms cos² ramps): an initial 500 ms noise segment followed by a pitch segment, i.e., T2_150,

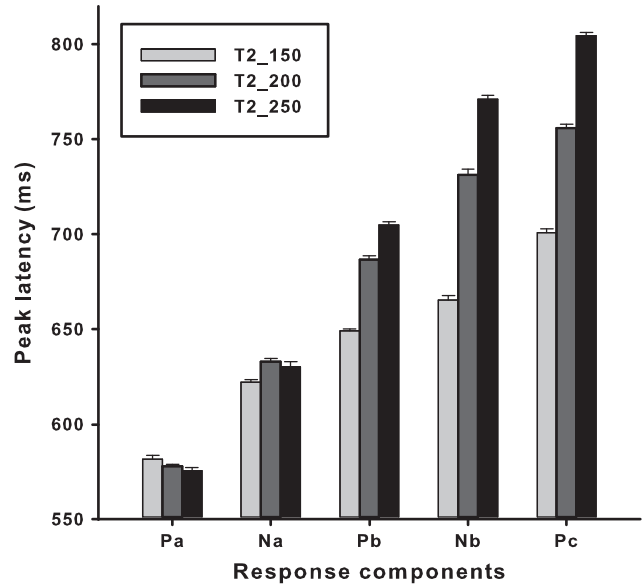


Fig. 4. Mean peak latency of Fz response components as a function of stimulus. From stimulus onset (Pa) to stimulus offset (Pc), peak latencies show increasing differentiation across stimuli (T2_150; T2_200; T2_250) reflecting sensitivity to increases in pitch acceleration and duration. No changes in peak latency are evident for Pa in contrast to marked changes across stimuli for Pb, Nb, and Pc. Na (pitch onset) reveals that the peak latency of T2_150 is shorter than either T2_200 or T2_250. Error bars do not appear because of the wide range of the Y-axis scale and the small size of SEs ($M=1.9$ ms).

T2_200, and T2_250 (Fig. 2). The overall RMS level of each segment was equated such that there was no discernible difference in intensity between initial and final segments. All stimuli were presented binaurally at 80 dB SPL through magnetically-shielded tubal insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA) with a fixed onset polarity (rarefaction) and a repetition rate of .94/s. Stimulus presentation order was randomized both within and across participants. All stimuli were generated and played out using an auditory evoked potential system (SmartEP, Intelligent Hearing Systems; Miami, FL, USA).

2.3. Cortical pitch response acquisition

Participants reclined comfortably in an electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts), ignore the sounds they heard, and were allowed to sleep throughout the duration of the recording procedure ($\approx 75\%$ fell asleep). The EEG was acquired continuously (5000 Hz sampling rate; .3–2500 Hz analog band-pass) using ASA-Lab EEG system (ANT Inc., The Netherlands) utilizing a 32-channel amplifier (REFA8-32, TMS International BV) and WaveGuard (ANT Inc., The Netherlands) electrode cap with 32-shielded sintered Ag/AgCl electrodes configured in the standard 10–20-montage system. The high sampling rate of 5 kHz was necessary to recover the brainstem frequency following responses in addition to the relatively slower cortical pitch components. Because the primary objective of this study was to characterize the cortical pitch components, the EEG acquisition electrode montage was limited to 9 electrode locations: Fpz, AFz, Fz, F3, F4, Cz, T7, T8, M1, M2 (Appendix C: figure, electrode montage). The AFz electrode served as the common ground and the common average of all connected unipolar electrode inputs served as default reference for the REFA8-32 amplifier. An additional bipolar channel with one electrode placed lateral to the outer canthi of the left eye and another electrode placed above the left eye was used to monitor artifacts introduced by ocular activity. Inter-electrode impedances were maintained below 10 k Ω . For each stimulus, EEGs were acquired in blocks of 1000 sweeps. The experimental protocol took about 2 h to complete.

2.4. Extraction of the cortical pitch response (CPR)

CPR responses were extracted off-line from the EEG files. To extract the cortical pitch response components, EEG files were first down sampled from 5000 Hz to 2048 Hz. They were then digitally high-pass filtered (3–25 Hz) to enhance the transient components and minimize the sustained component. Sweeps containing electrical activity exceeding $\pm 50 \mu\text{V}$ were rejected automatically. Subsequently, averaging was performed on all 8 unipolar electrode locations using the common reference to allow comparison of CPR components at the right frontal (F4), left

frontal (F3), right temporal (T8), and left temporal (T7) electrode sites to evaluate laterality effects. The re-referenced electrode site, Fz-linked T7/T8, was used to characterize the transient pitch response components. 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). This identical electrode configuration will also allow us to compare these CPR responses with brainstem responses in subsequent experiments. For both averaging procedures, the analysis epoch was 1200 ms including the 100 ms pre-stimulus baseline.

2.5. Analysis of CPR

The CPR is characterized by obligatory components (P1/N1) corresponding to the onset of energy in the precursor noise segment of the stimulus followed by several transient response components (Pa, Na, Pb, Nb, Pc, Nc) occurring after the onset of the pitch-eliciting segment of the stimulus. Because our primary objective in this study is to characterize what aspects of the dynamic pitch contours are being indexed by the components of the CPR (e.g., pitch onset, pitch acceleration, stimulus duration), only the latency and magnitude of the CPR components were evaluated. Both peak latencies (time interval between pitch-eliciting stimulus onset and a response peak of interest) of response components Pa, Na, Pb, Nb, and Pc, and interpeak latencies (time interval between response peaks) Pa–Na, Na–Pb, Pb–Nb, Nb–Pc, and Pa–Nc were measured to enable us to identify the components associated with pitch onset, pitch acceleration, stimulus duration, and stimulus offset. Peak-to-peak amplitude of Pa–Na, Na–Pb, Pb–Nb, and Nb–Pc was measured to determine if variations in amplitude were indexing specific aspects of the pitch contour (e.g., pitch acceleration). In addition, peak-to-peak amplitude of Na–Pb and Pb–Nb was measured separately at the frontal (F3/F4) and temporal (T7/T8) electrode sites to evaluate laterality effects. To enhance visualization of the laterality effects along a spectrotemporal dimension, joint time–frequency analysis (using a continuous wavelet transform) was performed on the grand average waveforms derived from the frontal and temporal electrodes. Since the focus of this paper is on the characterization of the CPR components, the obligatory onset responses to the noise precursor were not analyzed.

2.6. Statistical analysis

Separate two-way, mixed model ANOVAs were conducted on peak latency, interpeak latency, and peak-to-peak amplitude derived from the Fz electrode site. Subjects were treated as a random factor, stimulus and component as within-subject factors. In the analysis of peak latency, there were five components (Pa, Na, Pb, Nb, Pc); in the analysis of interpeak latency and peak-to-peak amplitude, four components (Pa–Na, Na–Pb, Pb–Nb, Nb–Pc). By analyzing these components, we were able to assess the effects of pitch acceleration on latency and amplitude across stimuli (T2_150, T2_200, T2_250). Separate three-way mixed model ANOVAs were conducted on peak-to-peak amplitude derived from the T7/T8 and F3/F4 electrode sites. Subjects were treated as a random factor; stimulus (T2_150, T2_200, T2_250), component (Na–Pb, Pb–Nb), and side (left, right) as within-subject factors. By focusing on these two components, we were able to determine whether laterality effects at the frontal and temporal sites vary as a function of stimulus.

3. Results

3.1. Response morphology of CPR components

Grand averaged cortical evoked response components to the three stimuli are shown in Fig. 3. The top panel shows both the superimposed P1/N1 onset complex for the three stimuli (black) and the CPR component to T2_250 (green). As expected (Krishnan, Bidelman et al., 2012), the obligatory P1/N1 complex, reflecting neural activity synchronized to the onset of the noise precursor (black), is very similar across the three stimulus conditions. The CPR components, elicited by the pitch-eliciting stimulus segment, are characterized by a series of successive biphasic components (in ms): Pa, 70–85; Na, 125–140; Pb, 700–715; Nb, 765–775; Pc, 800–815; and Nc, 840–855, for the T2_250 stimulus only. Na, Pb, and Nb were the most robust response components. The bottom panel in Fig. 3 shows only the CPR waveforms elicited by the three stimuli. It is clear from these waveform traces that peak latency for Pa and Na do not change appreciably across stimulus conditions (alignment of latency indicated by solid vertical lines). In contrast, response components Pb, Nb, Pc, and Nc all show a systematic

increase in peak latency across stimulus conditions (indicated by dashed lines). Consistent with these observations, the interpeak latencies (Na–Pb, Pb–Nb, Nb–Pc) increase across stimulus conditions. Response amplitude for Na, Pb, and Nb also appears to increase across stimuli with T2_150 showing the smallest amplitude. These systematic changes in response latency and amplitude are likely produced by the decrease in rate of pitch acceleration and the increase in stimulus duration across the three stimulus conditions.

3.2. Latency of CPR components

The stimulus effects on mean peak latency of the CPR components are displayed in Fig. 4 (Appendix D1: table, *M*, *SE*). Qualitatively, we observe large increases in peak latency across stimuli for components Pb, Nb, and Pc, with smaller increases in Na. An omnibus two-way ANOVA yielded significant ($p < .001$) main effects (stimulus, $F_{2,18} = 1029.61$; component, $F_{4,45} = 3116.71$), and an interaction between stimulus and component ($F_{8,72} = 213.42$). By Pb, Nb, and Pc separately, post hoc multiple comparisons of stimuli ($\alpha_{Bonferroni} = .05$) revealed that peak latency of T2_250 was longer than T2_200 which, in turn, was longer than T2_150. For Na, T2_150 was shorter than either of the other two stimuli. For Pa, differences in peak latency between stimuli failed to reach significance. These results suggest that decreasing acceleration rate and increasing stimulus duration increases peak latency of components Pb, Nb, and Pc, with a smaller change for Na, and no change for Pa. The invariant nature of Pa latency across stimuli suggests that it may be indexing the onset of the pitch-eliciting stimulus. The latency of Na is consistent with a pitch-onset response component. In contrast, the relatively large changes in latency for Pb, Nb, and Pc across stimuli are likely to represent changes in the more dynamic portions of the pitch contour and stimulus duration.

Fig. 5 displays mean interpeak latency effects of stimuli on pitch-related components: Pa–Na, Na–Pb, Pb–Nb, Nb–Pc (Appendix D2: table, *M*, *SE*). Na–Pb and Pb–Nb are strikingly similar in showing a systematic increase in their response patterns across stimuli. An omnibus two-way ANOVA on the pitch-related components yielded significant ($p < .001$) main effects (stimulus, $F_{2,72} = 131.13$; component, $F_{3,36} = 40.09$), and an interaction between stimulus and component ($F_{6,72} = 31.33$). By Na–Pb and Pb–Nb separately, post hoc paired comparisons of stimuli indicated that interpeak latency of T2_250 was longer than T2_200 which, in turn, was longer than

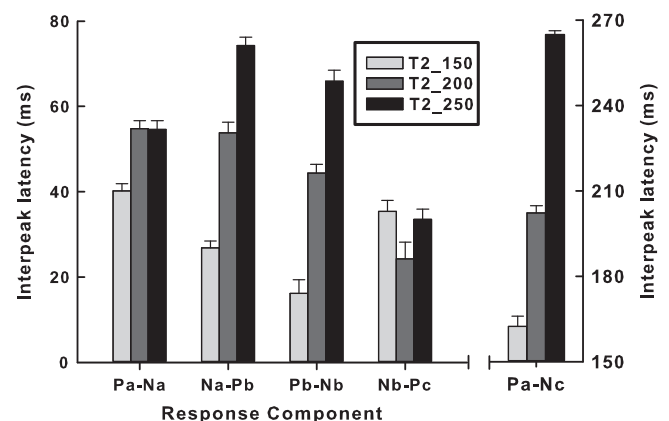


Fig. 5. Mean interpeak latency of Fz response components as a function of stimulus. Two pitch-relevant components (Na–Pb, Pb–Nb) are strikingly similar with proportionate increases in interpeak latency from T2_150 to T2_250 (left four response components). Pa–Na (sound onset to pitch onset) reveals that the interpeak latency of T2_150 is shorter than either T2_200 or T2_250. In the case of Nb–Pc, T2_200 is shorter than either T2_150 or T2_250. Pa–Nc is most accurate in tracking the onset and offset of the stimulus (rightmost response component). Error bars = \pm 1SE.

T2_150. For Pa–Na, T2_150 was shorter than T2_200 and T2_250. For Nb–Pc, the pattern was somewhat irregular; T2_200 was shorter than T2_250 and T2_150. An omnibus two-way ANOVA yielded significant ($p < .001$) main effects (stimulus, $F_{2,18} = 1029.61$; component, $F_{4,45} = 3116.71$), and a significant interaction between stimulus and component ($F_{8,72} = 213.42$). The large and systematic changes across stimuli for both Na–Pb and Pb–Nb suggest that they may be indexing the more rapidly-changing portions of the pitch contours.

To determine which response temporal interval best reflected stimulus duration, we conducted a two-way mixed model ANOVA on the interpeak latency of two response components (Pa–Pc, Pa–Nc) and three stimuli (T2_150, T2_200, T2_250). Results yielded significant ($p < .001$) main effects (stimulus, $F_{2,18} = 637.33$; component, $F_{1,18} = 360.88$), and an interaction between stimulus and component ($F_{2,18} = 25.44$). Pa–Nc exhibited longer interpeak latency than Pa–Pc, and corresponded best with stimulus duration irrespective of stimulus (Fig. 5; Appendix E: figure, Pa–Nc vs. Pa–Pc). The interaction effect was due to a larger relative difference between Pa–Nc and Pa–Pc for T2_150 (+1.36) than for either T2_200 (+1.13) or T2_250 (+1.16).

To confirm that Pc and Nc are indeed stimulus offset components (and not related to pitch), we recorded responses from five of the participants to a “noise-to-noise” stimulus. It was comprised of the same noise precursor as T2_250, but the pitch segment was replaced by a 250 ms noise segment to eliminate pitch-evoked components. A comparison of the responses to the noise-to-pitch and noise-to-noise stimuli is displayed in Fig. 6. It clearly shows that responses to the noise-to-noise stimulus contained only Pc and Nc components, and moreover, that they overlapped temporally with the Pc and Nc components elicited by the noise-to-pitch stimulus. Thus it is clear that both Pc and Nc components reflect stimulus offset, and the Pa–Nc interval better approximates stimulus duration. The mean Nc latency observed herein is also consistent with offset latencies previously reported for the cortical pitch response (Gutschalk, Patterson, Scherg, Uppenkamp, & Rupp, 2007; Seither-Preisler, Krumbholz, Patterson, Seither, & Lutkenhoner, 2004; Seither-Preisler et al., 2006).

3.3. Amplitude of CPR components

The stimulus effects on peak-to-peak amplitude of the CPR components are displayed in Fig. 7 (Appendix D3: table; *M*, *SE*). Again, Na–Pb and Pb–Nb are strikingly similar in their response patterns across stimuli. An omnibus two-way ANOVA on the pitch-

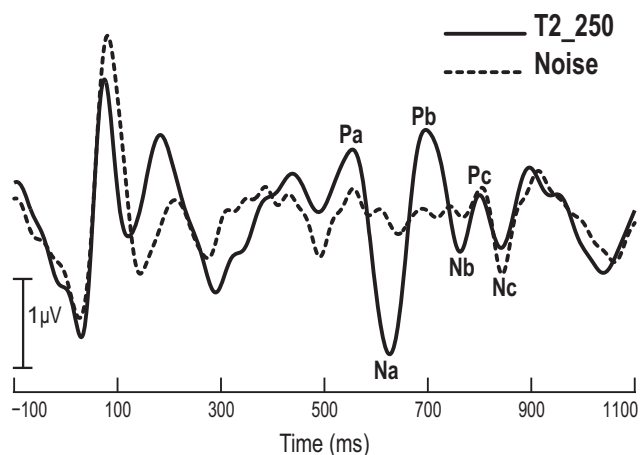


Fig. 6. Grand averaged cortical evoked response components at the Fz electrode site for the pitch condition (T2_250, solid line) and the noise condition (dashed line). A comparison of the two waveforms shows that components Pc and Nc are aligned in time, indicating that both of them are indexing stimulus offset.

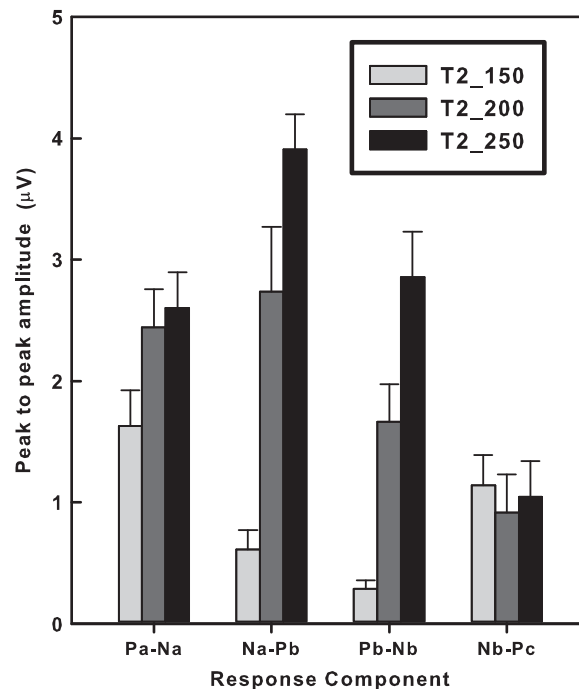


Fig. 7. Mean peak-to-peak amplitude of Fz response components as a function of stimulus. Two pitch-relevant components (Na–Pb, Pb–Nb) show that peak-to-peak amplitude increases steadily across the three stimuli ($T2_{150} < T2_{200} < T2_{250}$). In the case of Pa–Na, T2_250 is marginally higher in peak-to-peak amplitude than T2_150. Nb–Pc shows no differences in peak-to-peak amplitude between stimuli. Error bars = ± 1 SE.

related components yielded significant ($p < .001$) main effects (stimulus, $F_{2,18} = 34.87$; component, $F_{3,27} = 10.76$), and an interaction between stimulus and component ($F_{6,54} = 8.75$). By Na–Pb and Pb–Nb separately, post hoc paired comparisons of stimuli indicated that peak-to-peak amplitude of T2_250 was greater than T2_200 which, in turn, was greater than T2_150. In the case of Nb–Pc, differences in peak-to-peak amplitude between stimuli failed to reach significance. Within Pa–Na, peak-to-peak amplitude was marginally higher in T2_250 than T2_150 ($p = .042$). The latency and amplitude changes specific to Na–Pb and Pb–Nb reinforce our view that these components reflect pitch-relevant neural activity that may be closely associated with the dynamic portions of the pitch contour.

To support this view, we measured the extent to which CPR components (Pa–Na, Na–Pb, Pb–Nb, Nb–Pc) were associated with pitch acceleration, Pearson correlations (r) were computed between interpeak latency and peak-to-peak amplitude with the average acceleration (Hz/s) from minimum to maximum across stimuli (T2_150, 559; T2_200, 419; T2_250, 336). Na–Pb and Pb–Nb showed a significantly stronger negative association with pitch acceleration relative to the other two components (Table 2). The negative correlation coefficient means that both interpeak latency and peak-to-peak amplitude increase with decreasing acceleration. Na–Pb and Pb–Nb similarly showed a stronger negative association compared to other two components even when peak-to-peak amplitude was correlated with instantaneous acceleration (Fig. 1).

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

Fig. 8 displays the grand average waveforms of F3/F4 and T7/T8 (two left columns) and their corresponding spectra (two right columns) for each of the three stimuli. The waveform data reveal that pitch-related components at frontal F3 (orange) and F4 (green)

electrode sites essentially overlap for all three stimuli. In contrast, these same components at the right temporal electrode (T8: red) appear to be larger compared to the left temporal electrode (T7: blue), particularly for T2_250. Similar trends are evident in the spectrotemporal representations of the pitch-related components (enclosed within vertical dashed lines, 500–900 ms). Note the large rightward lateralization for T2_250.

Fig. 9 shows that hemispheric laterality effects on peak-to-peak amplitude of pitch-related components (Na–Pb, Pb–Nb) vary depending on electrode site (Fig. 9). These effects were evaluated by performing separate, three-way mixed model ANOVAs (component \times hemisphere \times stimulus) on peak-to-peak amplitude at the frontal (F3/F4) and temporal (T7/T8) electrode sites. For F3/F4, results revealed significant main effects of component ($F_{1,36}=11.52, p=.0017$) and stimulus ($F_{2,36}=44.03, p<.0001$). Other effects failed to reach significance including the interaction between stimulus and hemisphere. For T7/T8, on the other hand,

results showed significant main effects (component: $F_{1,34}=8.50, p=.0062$; stimulus: $F_{2,34}=43.55, p<.0001$; hemisphere: $F_{2,36}=14.78, p=.0005$) but also, and more importantly, an interaction between stimulus and hemisphere ($F_{2,34}=9.16, p=.0007$). By stimulus, post hoc multiple comparisons ($\alpha_{Bonferroni}=.05$) showed that a RH advantage (T8 > T7) was evident for T2_250 only. By hemisphere, all pairwise comparisons were significant (T2_250 > T2_200, T2_250 > T2_150, T2_200 > T2_150) except for T2_200 vs. T2_250 in the LH. These combined, stimulus- and hemisphere-dependent simple main effects suggest that the hemispheres are differentially sensitive to graded approximations to a prototypical pitch representation of Mandarin Tone 2. The main effect of component was the same for both F3/F4 and T7/T8. Na–Pb exhibited greater peak-to-peak amplitude than Pb–Nb across the three stimuli.

4. Discussion

Our objective is to identify cortical, pitch-specific responses that are elicited by linguistically-relevant, dynamic curvilinear pitch contours exemplary of those that occur in natural speech. Herein we examine the sensitivity of the multiple transient components of the CPR and relate them to specific temporal attributes associated with three, within-category variants of the Mandarin Tone 2. Our major findings show that the pitch onset component (Na) was invariant to changes in acceleration. In contrast, two components (Na–Pb, Pb–Nb) showed a strong correlation with pitch acceleration and a systematic increase in the interpeak latency and decrease in amplitude with increase in pitch

Table 2

Correlations between Fz components and pitch acceleration.

Component	Interpeak latency	Peak to peak amplitude ^a	Peak to peak amplitude ^b
Pa–Na	-.407 (.0255)	-.696 (<.0001)	-.411 (.0240)
Na–Pb	-.778 (<.0001)	-.954 (<.0001)	-.780 (<.0001)
Pb–Nb	-.777 (<.0001)	-.932 (<.0001)	-.776 (<.0001)
Nb–Pc	.053 (.7797)	.056 (.5621)	.056 (.7671)

Note: Values in parentheses represent levels of significance.

^a Average acceleration from minimum to maximum.

^b Instantaneous acceleration at 80 ms after pitch onset.

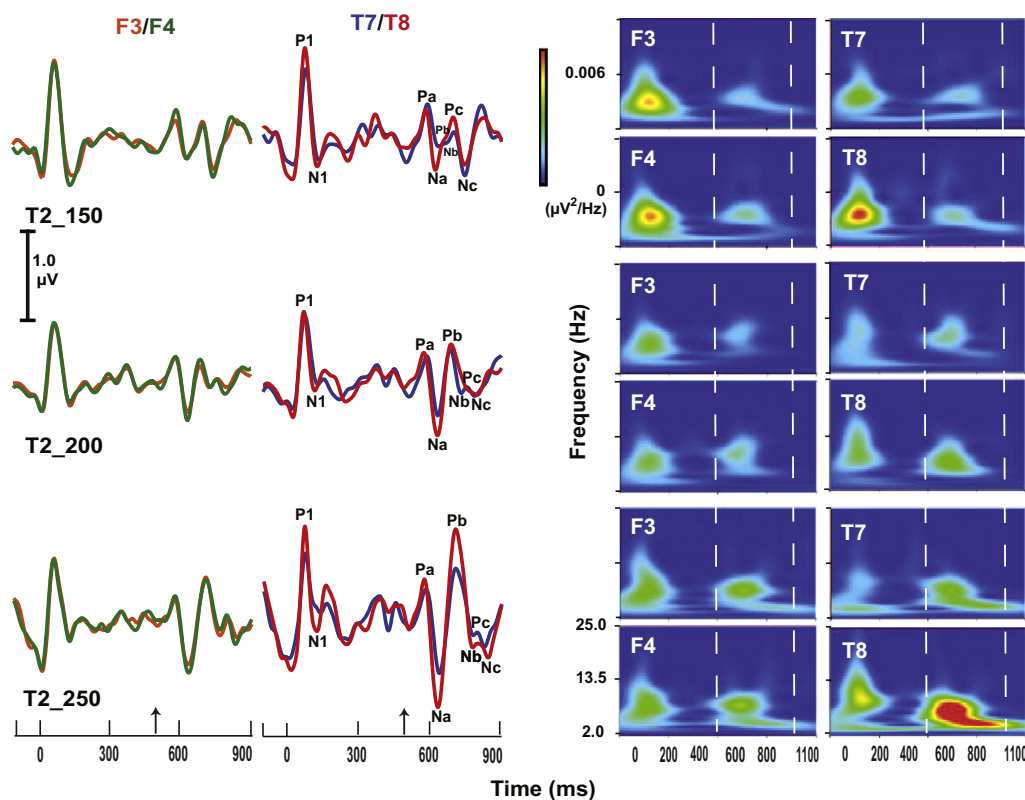


Fig. 8. Grand averaged cortical evoked response waveforms (left side) for the frontal F3 (orange) and F4 (green) and for the temporal T7 (blue) and T8 (red) electrode sites per stimulus condition. The upward arrow on the time axes denotes the onset of the pitch segment. The P1/N1 onset complex and the CPR components (Pa, Na, Pb, Nb, Pc, Nc) are identified. Responses at the frontal sites (F3, F4) essentially overlap; responses at the temporal sites, however, show larger responses at T8 compared to T7 for both T2_200 and T2_250. Spectrograms (wavelet analysis) at frontal and temporal electrode sites (right side) reveal the robust rightward asymmetry in spectrotemporal activity in response to T2_250 within the Na–Pb and Pb–Nb time windows (demarcated by two vertical dashed lines). See also caption to Fig. 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

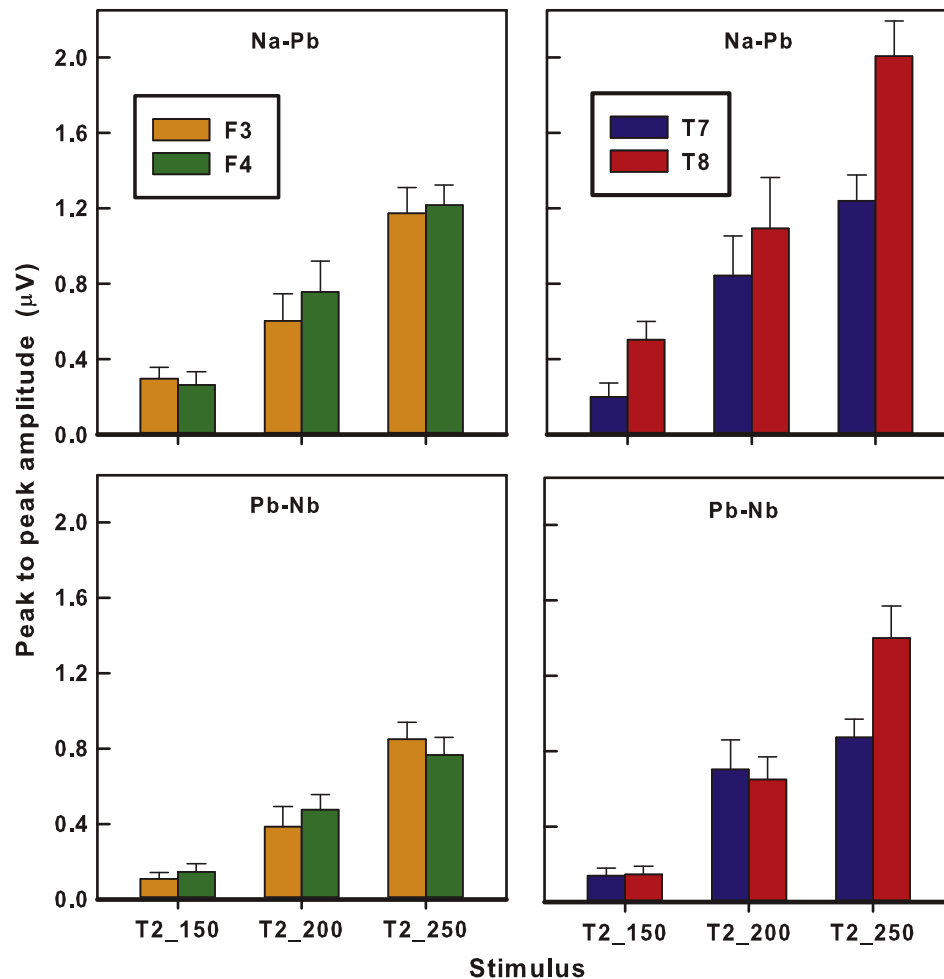


Fig. 9. Mean peak-to-peak amplitude of frontal F3/F4 (left column) and temporal T7/T8 (right column) response components (Na-Pb: top row; Pb-Nb: bottom row) as a function of stimulus. At the temporal sites, a right-sided preference is evident for T2_250 only. At the frontal sites, there is no evidence of a hemispheric preference for any stimulus. A gradual increase in amplitude is observed between stimuli at the right temporal site (T2_150 < T2_200 < T2_250). A similar increase in amplitude between stimuli is observed at the left temporal site (blue) except for T2_200/T2_250. LH, left hemisphere; RH, right hemisphere. Error bars = ± 1 SE. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acceleration that followed the time course of pitch change across the three stimuli. Pc–Nc marks unambiguously the stimulus offset. These data lead us to propose that in the early stages of cortical sensory processing of pitch, a series of neural markers flag different temporal attributes of a dynamic pitch contour: onset of temporal regularity (Na); changes in temporal regularity between onset and offset (Na–Pb, Pb–Nb); and offset of temporal regularity (Pc–Nc). A rightward asymmetry at a temporal site is observed for the tonal variant (T2_250) that most closely approximates a prototypical representation of Tone 2. These findings highlight the emergence of early hemispheric preferences and their functional roles as related to sensory and cognitive properties of the stimulus.

4.1. Pitch-relevant neural components of the CPR

4.1.1. Pitch onset (Na)

In this study the invariance of Na latency and amplitude across high-iteration-step IRN stimuli suggests that the Na component may be specific to pitch onset and its salience, and moreover, that it may not be sensitive to other varying attributes of the stimuli, viz., duration or pitch acceleration. Evidence adduced from earlier studies point to the lateral Heschl's gyrus as a source for the pitch onset response component including dipole source analysis of the MEG pitch onset response (Chait et al., 2006; Gutschalk et al., 2004;

Krumbholz et al., 2003; Ritter et al., 2005; Seither-Preisler et al., 2004) and localization of direct depth electrode recordings along the lateral to medial stretch of Heschl's gyrus in humans (Schonwiesner & Zatorre, 2008). Previous studies using IRN stimuli show that the latency and amplitude of the pitch onset response varies systematically with the pitch salience (EEG: Krishnan, Bidelman et al., 2012; MEG: Krumbholz et al., 2003; Seither-Preisler et al., 2006; Soeta, Nakagawa, & Tonoike, 2005). Functional neuroimaging studies in humans (Griffiths et al., 1998; Griffiths, Uppenkamp, Johnsrude, Josephs, & Patterson, 2001) and intracranial electrode recordings in both primates (Bendor & Wang, 2005) and humans (Schonwiesner & Zatorre, 2008) demonstrate that activity of the primary auditory cortex increases as a function of the number of IRN iteration steps. 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 and/or neural synchrony in cortical auditory neurons in the lateral Heschl's gyrus. This inference about the CPR source, however, must await empirical validation by neuroimaging techniques with superior spatial resolution (e.g., MEG, fMRI).

4.1.2. Dynamic portions of pitch contour (Na–pb, pb–nb)

In contrast to the onset (Na) and offset (Pc–Nc) components, Na–Pb and Pb–Nb show a systematic increase in the interpeak

latency and decrease in amplitude with increase in pitch acceleration that parallels the time course of pitch change across the three stimuli. This finding suggests that these components may be indexing pitch-relevant neural activity associated with the more rapidly-changing portions of the pitch contour. The observation of a strong correlation with pitch acceleration for only these two components further reinforces our view that these components are indexing the dynamic portions of the pitch contour. While our current experimental design does not permit us to determine whether components Na–Pb and Pb–Nb are indexing different portions of the dynamic segment of the pitch contour, we hypothesize that Na–Pb (relatively longer latency and larger amplitude) indexes the increasing pitch acceleration between the turning point and the point of maximum acceleration in the stimulus; whereas Pb–Nb (shorter latency and smaller amplitude) indexes the shorter pitch deceleration segment between maximum acceleration and stimulus offset.

To reinforce our view that Na–Pb and Pb–Nb reflect neural activity associated with pitch acceleration, we need to rule out the possibility that changes in stimulus duration may have contributed to changes in latency and amplitude of the pitch-relevant information. Indeed, it is well known that pitch information is integrated over duration. Notwithstanding, it is unlikely that changes in latency and amplitude of Na–Pb and Pb–Nb can be attributed to differences in temporal integration for pitch. Compared to psychoacoustic data, our stimulus durations fall well outside the range of temporal integration effects on pitch (≈ 80 ms) and its salience for stimuli with resolved harmonics (Plack et al., 1995, 2011; Plack & White, 2000; White & Plack, 1998, 2003). Data from electrophysiological studies indicate that the amplitude of the N1(m) response increases as a function of stimulus duration until reaching an asymptotic level with a temporal integration window between 20 and 50 ms (Alain, Woods, & Covarrubias, 1997; Gage & Roberts, 2000; Joutsiniemi, Hari, & Vilkmann, 1989; Krumbholz et al., 2003; Onishi & Davis, 1968; Ostroff, McDonald, Schneider, & Alain, 2003; Ross et al., 2009). Using the cortical MEG component N1(m), a similar temporal integration window (30–50 ms) is found for the periodicity of speech sounds (Yrttiaho, Tiitinen, Alku, Miettinen, & May, 2010). Taken together, it is unlikely that changes in latency and amplitude of Na–Pb and Pb–Nb are confounded by temporal integration for pitch.

While it is generally agreed that lateral Heschl's gyrus is the putative source for the pitch onset response (Na), the generator sources for the remaining transient, pitch-relevant components (Pb, Nb) are unknown, and cannot be determined from this study. From a hierarchical processing perspective, it is nevertheless plausible that these later components (Na–Pb, Pb–Nb) may reflect neural activity from spatially distinct generators that represent later stages of processing, relative to Na, along a pitch processing hierarchy. Using IRN stimuli to evaluate pitch changes over the time course of a sound with fMRI (Patterson et al., 2002), they report that only when pitch changes are varied over a time course to produce a musical melody does pitch processing move beyond auditory cortex with asymmetries emerging that favor the RH. In their experiment, the time course spans a sequence of discrete changes in flat pitch that constitute a musical unit. In the case of tonal languages, the time course spans a sequence of continuous changes in dynamic pitch within a single syllable that represent a linguistic unit.

4.1.3. Stimulus offset (Pc–Nc)

Our findings clearly show that the biphasic component Pc–Nc marks unambiguously the stimulus offset (Gutschalk et al., 2007; Hari et al. 1987; Seither-Preisler et al., 2004, 2006). However, it has also been suggested that rather than a simple offset of sound

energy, the offset response may in certain conditions represent cessation of more specific sound features (Seither-Preisler et al., 2006; e.g., pitch) or persistence of the response to stimulus regularity (Lutkenhoner, Seither-Preisler, Krumbholz, & Patterson, 2011). In this study the absence of significant changes in either amplitude or latency of the offset response across stimuli does not support the view that offset components may also carry pitch information. To the contrary, we observe essentially identical offset components both in latency and amplitude for the noise-to-pitch vs. noise-to-noise stimulus condition (Fig. 6). This finding strengthens our view that the offset response observed in this study is specific to stimulus (sound) offset and distinct from components that carry pitch-relevant information.

Viewing our CPR components as temporal regularity monitors (Lutkenhoner et al., 2011), we propose that the Na component marks the onset of stimulus regularity; the Pc–Nc component marks cessation of stimulus regularity (not pitch). The intervening transient components (Na–Pb, Pb–Nb) reflect additional temporal monitors in the time course of this stimulus regularity processing that may be indexing attributes of pitch change presented in a dynamic pitch contour. Thus, we propose a series of temporal regularity monitors that mark different temporal attributes of a dynamic pitch contour: (1) onset of temporal regularity; (2) changes in temporal regularity between onset and offset; and (3) offset of temporal regularity. It has already been proposed that temporal regularity monitors at onset and offset of stimulus regularity are embedded in the early stages of cortical processing (Lutkenhoner et al., 2011). By extension, we similarly propose that the temporal regularity monitor(s) that indexes attributes of dynamic pitch between the onset and offset are also embedded in the early stages of cortical sensory processing.

4.2. Hemispheric preferences for pitch processing

We first observe that a rightward asymmetry is confined to the temporal electrodes (T8 > T7) only. This asymmetry in the temporal lobe stands in stark contrast to the absence of hemispheric asymmetry in the frontal electrodes (F3/F4). Important to note is that our experimental protocol is free of task demands; stimuli are reduced to the pitch parameter only; and moreover, this differential pattern of hemispheric asymmetry is based on peak-to-peak amplitude responses extracted from the two putative, pitch-specific components (Na–Pb, Pb–Nb). Thus, it seems reasonable to infer that the temporal asymmetry reflects selective recruitment of pitch-specific mechanisms in right auditory cortex at an early stage of cortical processing. This finding converges with an extensive literature that attests to the greater role of the RH in the processing of pitch (Friederici & Alter, 2004; Meyer, 2008; Poeppel, Idsardi, & van Wassenhove, 2008; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; Zatorre & Gandour, 2008, for reviews).

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

Even though all three stimuli represent variants of a single tonal category, we observe differences in peak-to-peak amplitude of Na–Pb and Pb–Nb related to an interaction between hemisphere and stimulus. By hemisphere, we observe a gradual, continuous

effect across stimuli in the RH (T8) as measured by ($T2_250 > T2_200 > T2_150$). In the LH (T7), we do not find a continuous effect across all stimuli. Like the RH, $T2_250$ and $T2_200$ are greater in amplitude than $T2_150$, but $T2_250$ does not differ from $T2_200$. By stimulus, only $T2_250$ elicits a preference for the RH ($T8 > T7$). The question now is whether simple acoustic features are sufficient to explain this pattern of hemispheric specialization. Differences in neural responses between the left and right auditory cortices have been conceptualized as related to differences in the speed with which dynamically changing spectral information is processed. One theoretical approach emphasizes different temporal integration windows (Poeppel, 2003, see also Section 4.1.2); the other, a relative trade-off in temporal and spectral resolution (Zatorre et al., 2002). Both models can account for the observed stimulus continuum in the RH (Fig. 9). The RH preferentially extracts information from long integration windows (Poeppel, 2003, 150–250 ms). Encoding of pitch information becomes progressively better the longer the sampling window, and as a consequence, the frequency representation (Zatorre et al., 2002). Only Zatorre et al.'s model, however, can explain why only $T2_250$ elicits a RH preference, whereas $T2_200$ and $T2_150$ do not. Because all three stimuli fall into Poeppel's long integration window, it is unclear why the RH advantage is observed for $T2_250$ only. Neither model can adequately account for the absence of a stimulus effect between $T2_200$ and $T2_250$ in the LH.

It is incontrovertible that neural specializations are based on low-level features of the stimulus. More recent data from the extant literature, however, suggests that they can also be influenced by their functional status within one's domain of pitch expertise (Gandour & Krishnan, 2014; Krishnan, Gandour et al., 2012; Zatorre & Baum, 2012; Zatorre & Gandour, 2008, for reviews). Indeed, a complete account of pitch processing must allow for interactions between sensory and cognitive contributions that interact *within the same time interval*, as well as at different time intervals at different cortical levels of the brain. In this study, the time interval occurs at an early, preattentive stage of cortical sensory processing; all three stimuli fall within the bounds of a single tonal category. The stimulus-dependent asymmetry at the right temporal site suggests that it is sensitive to graded approximations to a prototypical pitch representation of Mandarin Tone 2. The stimulus with the most gradual change in pitch acceleration, $T2_250$, appears to cross that threshold that leads to a RH advantage at this temporal site ($T8 > T7$). Yet within the LH, $T2_250$ is indistinguishable from $T2_200$. By assuming that the LH is preferentially engaged for mediating the categorical status of pitch, this finding suggests that $T2_200$ as well as $T2_250$ are better candidates as representations of Mandarin Tone 2. Thus, within this early cortical time window we begin to see the emergence of hemispheric asymmetries and how the complementary roles of the two hemispheres reflect influences from sensory and cognitive properties of the stimulus.

Our findings are consistent with earlier ERP data that reveal the emergence of experience-dependent asymmetries at early cortical levels of processing. For example, in response to pure tones, musicians exhibit larger amplitudes of the early cortical components (N19m–P30m) within the right primary auditory cortex of trained musicians only; non-musicians show no hemispheric asymmetry (Schneider et al., 2002). In response to musical stimuli, a right temporal advantage is seen in the cortical N1 component related to pitch transition (change-N1, ~100 ms latency) in trained musicians (Itoh, Okumiya-Kanke, Nakayama, Kwee, & Nakada, 2012). But no hemispheric asymmetry is observed for the onset component. Both aforementioned studies point to experience-dependent enhancement of pitch processing in the right auditory cortex related to the music domain. Both employ discrete, flat pitch

stimuli that are of musical relevance. This study uses dynamic, curvilinear pitch stimuli that are relevant to the language domain. What is especially interesting is the discovery of ERP components related to *changes in pitch* regardless of domain (cf. Itoh et al., 2012). We acknowledge that a direct comparison of data from English listeners for the same three tonal stimuli must be completed in order to confirm an experience-dependent effect on our pitch-related components (in progress).

4.3. Conclusions

Our discovery of ERP components sensitive to specific attributes of dynamic, curvilinear pitch contours that are ecologically representative of natural speech opens up a new avenue for research into stages of pitch processing in the human brain. The experimental paradigm used in this study may also provide a new window to evaluate the online interplay 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), and how experience shapes pitch processing at cortical and subcortical levels. The selection of dynamic pitch contours representative of lexical tone will enable us to evaluate the influence of language experience on the latency and amplitude of these cortical pitch response components in subsequent experiments. Complementary studies using MEG will be crucial to determine the anatomical sources of these components in an effort to shed more light on specific cortical generators contributing to the hierarchical stages of pitch processing, and how experience may shape these processes.

Acknowledgments

Research supported by NIH 5R01DC008549-06 (A.K.). Thanks to Jingyi Zhu and Longjie Cheng for their assistance with statistical analysis (Department of Statistics); Jilian Wendel and Chandan Hunsur Suresh for their help with data acquisition and graphics, respectively. Reprint requests should be addressed to Ananthanarayan Krishnan, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN 47907-2038, USA, or via email: rkrish@purdue.edu.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.neuropsychologia.2014.04.006>.

References

- Alain, C., Woods, D. L., & Covarrubias, D. (1997). Activation of duration-sensitive auditory cortical fields in humans. *Electroencephalography and Clinical Neurophysiology*, 104(6), 531–539.
- Alku, P., Sivonen, P., Palomaki, K., & Tiitinen, H. (2001). The periodic structure of vowel sounds is reflected in human electromagnetic brain responses. *Neuroscience Letters*, 298(1), 25–28.
- Barker, D., Plack, C. J., & Hall, D. A. (2012). Reexamining the evidence for a pitch-sensitive region: A human fMRI study using iterated ripple noise. *Cerebral Cortex*, 22(4), 745–753, <http://dx.doi.org/10.1093/cercor/bhr065>.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436(7054), 1161–1165, <http://dx.doi.org/10.1038/nature03867>.
- Besson, M., Chobert, J., & Marie, C. (2011). Language and music in the musician brain. *Language and Linguistics Compass*, 5(9), 617–634, <http://dx.doi.org/10.1111/j.1749-818x.2011.00302>.
- Cariani, P. A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology*, 76(3), 1698–1716.

- Cariani, P. A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *Journal of Neurophysiology*, 76(3), 1717–1734.
- Cedolin, L., & Delgutte, B. (2005). Pitch of complex tones: Rate-place and interspike interval representations in the auditory nerve. *Journal of Neurophysiology*, 94(1), 347–362. <http://dx.doi.org/10.1152/jn.01114.2004> (doi: 01114.2004 [pii]).
- Chait, M., Poeppel, D., & Simon, J. Z. (2006). Neural response correlates of detection of monaurally and binaurally created pitches in humans. *Cerebral Cortex* (New York, NY: 1991), 16(6), 835–848. doi: bhj027 [pii] 10.1093/cercor/bhj027.
- Chandrasekaran, B., Gandour, J. T., & Krishnan, A. (2007). Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. *Restorative Neurology and Neuroscience*, 25(3–4), 195–210.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, 1128(1), 148–156. <http://dx.doi.org/10.1016/j.brainres.2006.10.064> (doi: S0006-8993(06)03179-9 [pii]).
- Denham, S. (2005). Pitch detection of dynamic iterated rippled noise by humans and a modified auditory model. *Biosystems*, 79(1–3), 199–206.
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16(5), 419–423.
- Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, 89(2), 267–276. [http://dx.doi.org/10.1016/S0093-934X\(03\)00351-1](http://dx.doi.org/10.1016/S0093-934X(03)00351-1) (S0093934 x 03003511 [pii]).
- Gage, N. M., & Roberts, T. P. (2000). Temporal integration: Reflections in the M100 of the auditory evoked field. *Neuroreport*, 11(12), 2723–2726.
- Gandour, J. T., & Krishnan, A. (2014). Neural bases of lexical tone. In: H. Winkler, & P. Padakannaya (Eds.), *Handbook of South and Southeast Asian psycholinguistics* (pp. 339–349). Cambridge, UK: Cambridge University Press.
- Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1(5), 422–427.
- Griffiths, T. D., Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, 20(12), 1128–1132. <http://dx.doi.org/10.1016/j.cub.2010.04.044>.
- Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, 4(6), 633–637.
- Gu, F., Zhang, C., Hu, A., Zhao, G., & Zhang, X. (2013). Left hemisphere lateralization for lexical and acoustic pitch processing in Cantonese speakers as revealed by mismatch negativity. *Neuroimage*. 10.1016/j.neuroimage.2013.02.080.
- Gutschalk, A., Patterson, R. D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage*, 15(1), 207–216. <http://dx.doi.org/10.1006/nimg.2001.0949>.
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *Neuroimage*, 22(2), 755–766. <http://dx.doi.org/10.1016/j.neuroimage.2004.01.025> (S1053811904000680 [pii]).
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2007). The effect of temporal context on the sustained pitch response in human auditory cortex. *Cerebral Cortex*, 17(3), 552–561. <http://dx.doi.org/10.1093/cercor/bhj180>.
- Hari, R., Pelizzzone, M., Makela, J. P., Hallstrom, J., Leinonen, L., & Louvasmaa, O. V. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology*, 26(1), 31–43.
- Itoh, K., Okumiyama-Kaneko, Y., Nakayama, Y., Kwee, I. L., & Nakada, T. (2012). Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *European Journal of Neuroscience*, 36(11), 3580–3592. <http://dx.doi.org/10.1111/j.1460-9568.2012.08278.x>.
- Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, 123(Part 1), 155–163.
- Joutsiniemi, S. L., Hari, R., & Vilkmann, V. (1989). Cerebral magnetic responses to noise bursts and pauses of different durations. *Audiology*, 28(6), 325–333.
- Koelsch, S. (2012). *Brain & music*. Chichester, UK: Wiley-Blackwell.
- Kratochvil, P. (1985). Variable norms of tones in Beijing prosody. *Cahiers de Linguistique Asie Orientale*, 14(2), 153–174.
- Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, 16(2), 105–110.
- Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., & Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, 50(12), 2849–2859. <http://dx.doi.org/10.1016/j.neuropsychologia.2012.08.013>.
- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2012). Experience-dependent plasticity in pitch encoding: From brainstem to auditory cortex. *Neuroreport*, 23(8), 498–502. <http://dx.doi.org/10.1097/WNR.0b013e328353764d>.
- Krishnan, A., Gandour, J. T., Smalt, C. J., & Bidelman, G. M. (2010). Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech. *Brain and Language*, 114(3), 193–198. <http://dx.doi.org/10.1016/j.bandl.2010.05.004>.
- Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, 21(6), 1092–1105. <http://dx.doi.org/10.1162/jocn.2009.21077>.
- Krumholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lutkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13(7), 765–772.
- Luo, H., Ni, J. T., Li, Z. H., Li, X. O., Zhang, D. R., Zeng, F. G., et al. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences of the United States of America*, 103(51), 19558–19563.
- Lutkenhoner, B., Seither-Preisler, A., Krumholz, K., & Patterson, R. D. (2011). Auditory cortex tracks the temporal regularity of sustained noisy sounds. *Hearing Research*, 272(1–2), 85–94. <http://dx.doi.org/10.1016/j.heares.2010.10.013>.
- Lutkenhoner, B., Seither-Preisler, A., & Seither, S. (2006). Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. *Neuroimage*, 30(3), 927–937. <http://dx.doi.org/10.1016/j.neuroimage.2005.10.034>.
- Lutkenhoner, B., & Steinstrater, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neuro-Otology*, 3, 191–213.
- Maess, B., Jacobsen, T., Schroger, E., & Friederici, A. D. (2007). Localizing pre-attentive auditory memory-based comparison: Magnetic mismatch negativity to pitch change. *Neuroimage*, 37(2), 561–571. <http://dx.doi.org/10.1016/j.neuroimage.2007.05.040> (doi: S1053-8119(07)00462-4 [pii]).
- Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211. <http://dx.doi.org/10.1162/089892906775783660>.
- Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, 23(10), 2701–2715. <http://dx.doi.org/10.1162/jocn.2010.21585>.
- Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. *Journal of the Acoustical Society of America*, 102(3), 1811–1820.
- Meyer, M. (2008). Functions of the left and right posterior temporal lobes during segmental and suprasegmental speech perception. *Zeitschrift für Neuropsychologie*, 19(2), 101–115.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102(3), 1864–1877.
- Munte, T. F., Altenmüller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews: Neuroscience*, 3(6), 473–478. <http://dx.doi.org/10.1038/nrn843>.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Onishi, S., & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked V potentials. *Journal of the Acoustical Society of America*, 44(2), 582–591.
- Ostroff, J. M., McDonald, K. L., Schneider, B. A., & Alain, C. (2003). Aging and the processing of sound duration in human auditory cortex. *Hearing Research*, 181(1–2), 1–7.
- Pantev, C., Ostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811–814.
- Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, 11(9), 369–372.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36(4), 767–776.
- Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience*, 24(30), 6810–6815. <http://dx.doi.org/10.1523/JNEUROSCI.0383-04.2004>24/30/6810 [pii].
- Plack, C. J., Carlyon, R. P., & Viemeister, N. F. (1995). Intensity discrimination under forward and backward masking: Role of referential coding. *Journal of the Acoustical Society of America*, 97(2), 1141–1149.
- Plack, C. J., Turgeon, M., Lancaster, S., Carlyon, R. P., & Gockel, H. E. (2011). Frequency discrimination duration effects for Huggins pitch and narrowband noise (L). *Journal of the Acoustical Society of America*, 129(1), 1–4. <http://dx.doi.org/10.1121/1.3518745>.
- Plack, C. J., & White, L. J. (2000). Perceived continuity and pitch perception. *Journal of the Acoustical Society of America*, 108(3 Part 1), 1162–1169.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, 41(1), 245–255.
- Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 363(1493), 1071–1086. <http://dx.doi.org/10.1098/rstb.2007.2160> (TM425571U1117682 [pii]).
- Ritter, S., Gunter Dosch, H., Specht, H. J., & Rupp, A. (2005). Neuro-magnetic responses reflect the temporal pitch change of regular interval sounds. *Neuroimage*, 27(3), 533–543. <http://dx.doi.org/10.1016/j.neuroimage.2005.05.003>.
- Ross, B., Snyder, J. S., Aalto, M., McDonald, K. L., Dyson, B. J., Schneider, B., et al. (2009). Neural encoding of sound duration persists in older adults. *Neuroimage*, 47(2), 678–687. <http://dx.doi.org/10.1016/j.neuroimage.2009.04.051>.
- Sayles, M., & Winter, I. M. (2007). The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Research*, 1171, 52–66. <http://dx.doi.org/10.1016/j.brainres.2007.06.098> (doi: S0006-8993(07)01509-0 [pii]).
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, 5(7), 688–694. <http://dx.doi.org/10.1038/nn871>.

- Schonwiesner, M., & Zatorre, R. J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, 187(1), 97–105, <http://dx.doi.org/10.1007/s00221-008-1286-z>.
- Seither-Preisler, A., Krumbholz, K., Patterson, R., Seither, S., & Lutkenhoner, B. (2004). Interaction between the neuromagnetic responses to sound energy onset and pitch onset suggests common generators. *European Journal of Neuroscience*, 19(11), 3073–3080, <http://dx.doi.org/10.1111/j.1460-9568.2004.03423.x>.
- Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). Evidence of pitch processing in the N100m component of the auditory evoked field. *Hearing Research*, 213(1–2), 88–98, <http://dx.doi.org/10.1016/j.heares.2006.01.003>.
- Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, 23(13), 5545–5552.
- Soeta, Y., & Nakagawa, S. (2008). The effects of pitch and pitch strength on an auditory-evoked N1m. *Neuroreport*, 19(7), 783–787, <http://dx.doi.org/10.1097/WNR.0b013e3282fe2085>.
- Soeta, Y., Nakagawa, S., & Tonoike, M. (2005). Auditory evoked magnetic fields in relation to iterated rippled noise. *Hearing Research*, 205(1–2), 256–261.
- Swaminathan, J., Krishnan, A., Gandour, J. T., & Xu, Y. (2008). Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Transactions on Biomedical Engineering*, 55(1), 281–287, <http://dx.doi.org/10.1109/TBME.2007.896592>.
- Tervaniemi, M., Kruck, S., De Baene, W., Schroger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *The European Journal of Neuroscience*, 30(8), 1636–1642, <http://dx.doi.org/10.1111/j.1460-9568.2009.06955.x>.
- Tsang, Y. K., Jia, S., Huang, J., & Chen, H. C. (2011). ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height. *Neuroscience Letters*, 487(3), 268–272, <http://dx.doi.org/10.1016/j.neulet.2010.10.035> (S0304-3940(10)01374-1 [pii]).
- Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Cortical encoding of pitch: Recent results and open questions. *Hearing Research*, 271(1–2), 74–87, <http://dx.doi.org/10.1016/j.heares.2010.04.015>.
- White, L. J., & Plack, C. J. (1998). Temporal processing of the pitch of complex tones. *Journal of the Acoustical Society of America*, 103(4), 2051–2063.
- White, L. J., & Plack, C. J. (2003). Factors affecting the duration effect in pitch perception for unresolved complex tones. *Journal of the Acoustical Society of America*, 114(6 Part 1), 3309–3316.
- Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, 156, 249–268, [http://dx.doi.org/10.1016/S0079-6123\(06\)56013-3](http://dx.doi.org/10.1016/S0079-6123(06)56013-3) (doi: S0079-6123(06)56013-3 [pii]).
- Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565–585.
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231, <http://dx.doi.org/10.1016/j.neuroscience.2010.06.077> (doi: S0306-4522(10)00949-8 [pii]).
- Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, 111(3), 1399–1413.
- Yrttiaho, S., Tiitinen, H., Alku, P., Miettinen, I., & May, P. J. (2010). Temporal integration of vowel periodicity in the auditory cortex. *Journal of the Acoustical Society of America*, 128(1), 224–234, <http://dx.doi.org/10.1121/1.3397622>.
- Yrttiaho, S., Tiitinen, H., May, P. J., Leino, S., & Alku, P. (2008). Cortical sensitivity to periodicity of speech sounds. *Journal of the Acoustical Society of America*, 123(4), 2191–2199, <http://dx.doi.org/10.1121/1.2888489>.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, 84(2), 566–572.
- Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, 10(7), e1001372, <http://dx.doi.org/10.1371/journal.pbio.1001372>.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6(1), 37–46.
- Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 363(1493), 1087–1104, <http://dx.doi.org/10.1098/rstb.2007.2161> (doi: J412P80575385013 [pii]).
- Zheng, H.-Y., Minett, J. W., Peng, G., & Wang, W. S.-Y. (2012). The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study. *Language and Cognitive Processes*, 27(2), 184–209, <http://dx.doi.org/10.1080/01690965.2010.520493>.