

Published in final edited form as:

Brain Lang. 2010 September ; 114(3): 193–198. doi:10.1016/j.bandl.2010.05.004.

Language-dependent pitch encoding advantage in the brainstem is not limited to acceleration rates that occur in natural speech

Ananthanarayan Krishnan¹, Jackson T. Gandour¹, Christopher J. Smalt², and Gavin M. Bidelman¹

¹Department of Speech Language Hearing Sciences, Purdue University, USA

²Department of Electrical and Computer Engineering, Purdue University, USA

Abstract

Experience-dependent enhancement of neural encoding of pitch in the auditory brainstem has been observed for only specific portions of native pitch contours exhibiting high rates of pitch acceleration, irrespective of speech or nonspeech contexts. This experiment allows us to determine whether this language-dependent advantage transfers to acceleration rates that extend beyond the pitch range of natural speech. Brainstem frequency following responses (FFRs) were recorded from Chinese and English participants in response to four, 250-ms dynamic click-train stimuli with different rates of pitch acceleration. The maximum pitch acceleration rates in a given stimulus ranged from low (0.3 Hz/ms; Mandarin Tone 2) to high (2.7 Hz/ms; 2 octaves). Pitch strength measurements were computed from the FFRs using autocorrelation algorithms with an analysis window centered at the point of maximum pitch acceleration in each stimulus. Between-group comparisons of pitch strength revealed that Chinese exhibit more robust pitch representation than English across all four acceleration rates. Regardless of language group, pitch strength was greater in response to acceleration rates within or proximal to natural speech relative to those beyond its range. Though both groups showed decreasing pitch strength with increasing acceleration rates, pitch representations of the Chinese group were more resistant to degradation. FFR spectral data were complementary across acceleration rates. These findings demonstrate that perceptually salient pitch cues associated with lexical tone influence brainstem pitch extraction not only in the speech domain, but also in auditory signals that clearly fall outside the range of dynamic pitch that a native listener is exposed to.

Keywords

auditory; human; brainstem; pitch; language; frequency following response (FFR); click trains; Mandarin Chinese; experience-dependent plasticity; speech perception

© 2010 Elsevier Inc. All rights reserved.

Author responsible for correspondence: Jackson T. Gandour, PhD Purdue University Department of Speech Language Hearing Sciences 1353 Heavilon Hall 500 Oval Drive West Lafayette, IN 47907-2038 USA Fax: 765.494.0771 Tel: 765.494.3821 gandour@purdue.edu. Reprint requests should be addressed to Ananthanarayan Krishnan, Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, IN, USA 47907-2038, or via rkrish@purdue.edu..

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

1. Introduction

There is an emerging literature to support the notion that the neural representation of pitch may be influenced by one's experience with language (or music) at subcortical as well as cortical levels of processing (for reviews, see Johnson, Nicol, & Kraus, 2005; Kraus & Banai, 2007; Kraus & Nicol, 2005; Krishnan & Gandour, 2009; Patel & Iversen, 2007; Tzounopoulos & Kraus, 2009; Zatorre & Gandour, 2008). Pitch provides an excellent window for studying language-dependent effects on subcortical processing as it is one of languages' most important information-bearing components. Tone languages are especially advantageous for investigating the linguistic use of pitch because variations in pitch patterns at the syllable level may be lexically significant (Yip, 2003). Mandarin Chinese, for example, has four lexical tones: *ma*¹ 'mother', *ma*² 'hemp', *ma*³ 'horse', *ma*⁴ 'scold' (Howie, 1976).

As a window into pitch processing in the brainstem, we measure electrophysiological activity using the human frequency-following response (FFR). This response reflects sustained phase-locked activity in a population of neural elements within the rostral brainstem (see Krishnan, 2006, for review). The FFR is characterized by a periodic waveform that follows the individual cycles of the stimulus waveform. FFRs can be elicited by a variety of stimuli that carry periodicity information, including low frequency tone bursts, complex tones, speech sounds, sinusoidal amplitude modulated sounds, frequency modulated sounds, click-trains, and iterated rippled noise (IRN). Experimental evidence overwhelmingly points to the inferior colliculus (IC) as the source of the FFR generator.

As reflected by FFRs, comparisons between native speakers of tone (Mandarin) and non-tone (English) languages show that native experience with lexical tones enhances pitch encoding at the level of the brainstem irrespective of speech or nonspeech context (Krishnan, Gandour, Bidelman, & Swaminathan, 2009; Krishnan, Xu, Gandour, & Cariani, 2005; Swaminathan, Krishnan, & Gandour, 2008). Language-dependent pitch encoding mechanisms in the brainstem are especially sensitive to the curvilinear shape of pitch contours that are exemplary of those that occur in natural speech. We fail to observe any language-dependent effects no matter how close a linearly accelerating or decelerating pitch pattern approximates a native lexical tone (Krishnan, Gandour, et al., 2009; Xu, Krishnan, & Gandour, 2006). Curvilinearity itself, though necessary, is insufficient to enhance pitch extraction of the auditory signal at the level of the brainstem. A nonnative curvilinear pitch pattern similarly fails to elicit a language-dependent effect (Krishnan, Gandour, et al., 2009). Using synthetic speech stimuli representative of Mandarin and Thai lexical tones, listeners from both tone languages are shown to be able to transfer their abilities in pitch encoding across languages (Krishnan, Gandour, & Bidelman, 2010b). Thus, brainstem neurons appear to be differentially sensitive to changes in specific dimensions of pitch without regard to their language identity as long as they occur in a language with a comparable phonological system. These findings collectively suggest that language-dependent neuroplasticity in the human brainstem occurs when dimensions of pitch in the auditory signal are part of the listener's experience and relevant to speech perception.

By analyzing pitch strength of individual sections of lexical tones, one of our major discoveries is that the degree of pitch acceleration or deceleration (i.e., rate of pitch change) is a critical dimension that influences pitch extraction in the brainstem. Chinese listeners exhibit more robust pitch representation than English primarily in those sections of lexical tones containing rapidly changing pitch regardless of the context, speech or nonspeech (Krishnan, Swaminathan, & Gandour, 2009). This heightened sensitivity to sections characterized by rapid changes in pitch is maintained even in severely degraded stimuli (Krishnan, Gandour, & Bidelman, 2010a). Thus, experience-dependent brainstem

mechanisms for pitch are especially sensitive to those dimensions of pitch contours that provide cues of high perceptual saliency in degraded as well as normal listening conditions.

Up to the present, all of our FFR experiments have employed pitch stimuli that fall within the bounds of a normal pitch range (male) for citation forms of lexical tones, and that exhibit rates of changes in pitch that occur in natural speech. The question then arises whether language-related expertise in pitch encoding of ecologically-relevant stimuli can transfer to pitch encoding of stimuli that exceed a normal pitch range, and are characterized by acceleration rates that do not occur in natural speech. Indeed, psychophysical studies of tone perception have demonstrated that the effects of linguistic experience may extend to nonspeech sounds under certain stimulus and task (discrimination, identification) conditions (Bent, Bradlow, & Wright, 2006; Luo, Boemio, Gordon, & Poeppel, 2007). Because of task confounds, however, behavioral tasks do not allow us to assess unambiguously to what extent the observed language-dependent effects are to be attributed to neurobiological properties of the auditory system. In our FFR experimental paradigm, we exploit a passive listening paradigm to index automatic, pitch encoding in the brainstem that involves no controlled memory or attention load, thus giving us a neurobiological index of language-dependent neuroplasticity at a subcortical, sensory level of processing.

The aim of the present study of pitch processing in the auditory brainstem was to determine the nature of language-dependent neuroplasticity in the processing of click-train homologues of a prototypical Mandarin Tone 2 (T2), and its variants thereof, along an acceleration-rate continuum that includes pitch contours that are not native to the Mandarin tonal space. The higher acceleration rates used in this study served to challenge the pitch extraction mechanism to temporally resolve and extract rapidly changing and therefore temporally degraded periodicity information beyond the range of natural speech. We hypothesized that pitch representation in response to these stimuli along the continuum would be more resistant to degradation in Chinese than English speakers regardless of their linguistic status. By using click trains, we preserved dynamic variations in pitch of auditory stimuli, but eliminated the formant structure characteristic of speech, thereby eliminating potential lexical semantic confounds. Click trains were employed instead of iterated rippled noise because they provide more stable periodicity information at higher acceleration rates (cf. Krishnan, Gandour, et al., 2009; Krishnan, Swaminathan, et al., 2009). FFRs were elicited in response to four click-train stimuli, varying along a continuum ranging from lower to higher rates of changes in pitch (Fig. 3). At one end of the continuum, the acceleration rate was that of an exemplary T2; at the other, the rate of acceleration extends well beyond the limits of natural speech, i.e., two octaves in 250 ms (cf. Xu & Sun, 2002). Of the two remaining pitch contours, one is of marginal relevance linguistically; the other exceeds the pitch range of T2 in citation form. Sensitivity to the rate of pitch change was indexed by pitch strength, and analyzed in a single analysis window that was centered at the time corresponding to the peak of the pitch acceleration curve in each stimulus. This maximum pitch acceleration window was chosen because language experience is observed to have an influence on pitch strength primarily in those tonal sections exhibiting higher degrees of acceleration or deceleration (Bidelman, Gandour, & Krishnan, in press; Krishnan, Swaminathan, et al., 2009; Swaminathan, Krishnan, & Gandour, 2008). In addition, correlation between stimulus and response spectra were examined to determine if representation of pitch relevant harmonics were more robust in the Chinese compared to the English group.

2. Results

2.1. Temporal and spectral properties of whole stimuli

Autocorrelograms (left panels) and narrow band spectrograms (right panels) derived from the grand averaged FFR waveforms in response to click-train f_0 contours representative of

Tone 2 (row 1), two intermediate rates of higher acceleration (rows 2-3), and an extraordinarily high acceleration rate (row 4) are shown in Fig. 1 for the Chinese and English groups (cf. Fig. 3). In the Chinese group, autocorrelograms show clear dark bands of phase-locked activity at f_0 and its multiples in response to an f_0 contour characterized by a linguistically-relevant acceleration rate (row 1), but less distinct and increasingly more diffuse bands in response to f_0 contours of greater acceleration (rows 2-4). In the English group, the bands are less distinct and more diffuse at all acceleration levels. At the highest acceleration rate, A4, the English group's correlogram appears to have no bands whatsoever in the region of interest (170 ms), unlike the Chinese group.

Consistent with the autocorrelograms, spectrograms in Fig.1 show more robust representation of the first 5 harmonics, particularly for the higher acceleration rates for the Chinese group. Stimulus-to-response spectral correlation coefficients within the region of maximum acceleration are reported for each of four acceleration rates per language group (Table 1). Results from an omnibus two-way (group \times acceleration) ANOVA on spectral correlation coefficients yielded significant main effects of group ($F_{1,18} = 20.57, p = 0.0003$) and acceleration ($F_{3,54} = 4.87, p < 0.0045$). The two-way interaction failed to reach significance ($F_{3,54} = 1.07, p = 0.3699$), meaning that the Chinese group response spectrum was more correlated with the stimulus spectrum as compared to the English across all acceleration rates, suggesting that the representation of pitch relevant harmonics were significantly more robust in the Chinese group for all stimuli.

2.2. Pitch strength of region of interest

FFR pitch strength of the region of maximum acceleration within the click-train stimuli is shown for each of four acceleration rates per language group (Fig. 2) (Supplementary Material: table; pitchstrength_M+SE.doc). Results from an omnibus two-way (group \times acceleration) ANOVA on pitch strength yielded significant main effects of group ($F_{1,18} = 19.11, p = 0.0004$) and acceleration ($F_{3,54} = 27.66, p < 0.0001$). The two-way interaction failed to reach significance ($F_{3,54} = 1.53, p = 0.2184$), meaning that pitch strength was greater in the native Chinese group as compared to the English across the board. Within both groups, post hoc Tukey multiple comparisons ($\alpha = 0.05$) revealed that pitch strength was greater at lower (A1, A2) than higher (A3, A4) acceleration rates. In the English group, pitch strength was also greater in A3 as compared to A4

3. Discussion

In this crosslanguage study, native speakers of a tone language, relative to those of a non-tone language, are observed to exhibit more robust pitch representation within tonal sections exhibiting maximal rates of pitch change regardless of their behavioral relevance. As measured by pitch strength, not only do the Chinese have an advantage over English listeners in response to an exemplar of Mandarin Tone 2, but also to scaled variants of T2 with increasingly higher maximum acceleration rates that fall proximal to or outside the boundary of natural speech. This finding demonstrates that neuroplasticity for pitch processing in the brainstem is not limited to the domain in which the pitch contours are behaviorally relevant (cf. Bidelman, et al., in press). Irrespective of language experience, there are substantial drops in sensitivity to those acceleration rates that extend well beyond the normal voice range (A1, A2 > A3, A4). Yet Chinese are still three times more sensitive than English in response to the pitch contour with the highest rate of pitch acceleration (A4).

3.1. Neural mechanisms underlying pitch encoding in the brainstem are differentially sensitive to increasing rates of pitch acceleration as a function of language experience

Our view is that crosslanguage differences in pitch representation reflect experience-dependent sensitivity of the neural mechanism underlying pitch encoding in native speakers of tone languages (Krishnan & Gandour, 2009). We have proposed that coincidence detection neurons in the inferior colliculus perform a correlation analysis (Langner, 1992, 2004) to extract pitch-relevant periodicities on the delayed and un-delayed temporal information arriving from the cochlear nucleus. This pitch information is then spatially mapped onto a periodicity pitch axis. We believe that long-term experience sharpens the tuning characteristics of the best modulation frequency neurons along the pitch axis with particular sensitivity to linguistically-relevant dynamic segments. This sharpening is likely mediated by local excitatory and inhibitory interactions that are known to play an important role in signal selection at the level of the brainstem (Ananthanarayan & Gerken, 1983, 1987). Such interaction may take the form of an active facilitation/disinhibition of the pitch intervals corresponding to the dynamic segments and inhibition of other pitch periods. Neuromodulatory inputs to the corticocollicular loop could influence the balance between excitation and inhibition (Xiong, Zhang, & Yan, 2009).

We observe herein a degradation of pitch representation with increasing pitch acceleration across language groups. This suggests a disruption in the correlation analysis to extract pitch-relevant periodicities. One plausible explanation for the disruption of this correlation analysis is that the coherence of the phase-locked neural activity representing pitch-relevant periodicities decreases with increasing acceleration of the fundamental frequency. This reduced coherence is due to temporal de-synchronization of the inputs to the pitch extraction mechanism. The fact that we also observe more robust pitch representation in the Chinese listeners even at higher acceleration rates suggests that their pitch extraction mechanism is less susceptible to de-synchronization. Both enhanced sensitivity and decreased susceptibility to degradation of pitch encoding for faster pitch acceleration rates in the Chinese listeners suggests an expansion of the dynamic range for this dimension of pitch. We argue that their heightened sensitivity to *rate-of-change of frequency*, a direct result of their language experience, is a crucial factor distinguishing tone language from non-tone language speakers/listeners.

Another finding is that multiple harmonics are better represented in Chinese than English listeners even at higher acceleration rates. This finding complements our data on voice pitch representation (Krishnan, et al., 2005). In the Chinese group, stronger pitch and more accurate pitch tracking co-occur with relatively stronger representation of pitch-relevant harmonics. Just the opposite is the case for the English group. Moreover, psychoacoustic and physiological data indicate that complex stimuli produce stronger and more accurate pitch percepts when spectral components are prominent in the dominance region for pitch (2nd to about the 5th harmonic) (Cariani & Delgutte, 1996; Schwartz & Purves, 2004).

3.2. Experience-dependent neuroplasticity in the brainstem is not circumscribed to the language domain

Background noise, competing sounds, or in this study, higher rates of pitch acceleration, represent a significant challenge for neural mechanisms encoding behaviorally relevant acoustic features of the target sound. The extent to which a particular *dimension* of pitch is resistant to such degraded listening conditions may serve as an index of its perceptual saliency. Regarding pitch acceleration, differential rates throughout the duration of a lexical tone are what determine its shape or *contour*. It is this pitch dimension that has been shown to be an important variable in separating tone from non-tone language speakers, as evidenced from perceptual judgments (Gandour, 1983), cortical evoked potentials

(Chandrasekaran, Gandour, & Krishnan, 2007), and, of course, subcortical responses in the brainstem (Krishnan, Swaminathan, et al., 2009). Despite their lack of pitch experience with these higher acceleration rates, Chinese listeners, as compared to English, are better able to transfer their abilities in pitch encoding from a *language domain* to a non-language or *auditory domain*. This finding suggests that brainstem neurons are differentially sensitive to changes in pitch even when presented with stimuli that do not occur within the listener's domain of experience.

4. Methods

4.1. Participants

Ten adult native speakers of Mandarin Chinese (6 male, 4 female), hereafter referred to as Chinese (C), and 10 adult monolingual native speakers of English without musical training (5 male, 5 female), referred to as English (E), participated in the FFR experiment. The three groups were closely matched in age (Chinese: $M = 23.8$, $SD = 1.8$; English: $M = 24.7$, $SD = 3.2$), years of formal education (Chinese: $M = 17.2$, $SD = 1.6$; English: $M = 18.1$, $SD = 3.6$). All were strongly right handed ($M = 87\%$) as measured by the Edinburgh Handedness inventory (Oldfield, 1971). All participants exhibited normal hearing sensitivity (better than 15 dB HL in both ears) at octave frequencies from 500 to 4000 Hz. Native speakers of Mandarin were born and raised in mainland China and none had received formal instruction in English before the age of 9 ($M = 11.7$, $SD = 1.1$). The English group had no prior experience learning a tonal language. Each participant completed a music history questionnaire (Wong & Perrachione, 2007). All participants had no more than 3 years of formal music training ($M = 0.7$, $SD = 0.8$) on any combination of instruments and none had any training within the past 5 years. All participants were paid and gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

4.2. Stimuli

A four-step stimulus continuum ranging from lower to higher rates of pitch acceleration was utilized to evoke brainstem responses (Fig. 3, top panel). One end of the pitch acceleration continuum was represented by the f_0 contour of Mandarin Tone 2 as produced in citation form by a male speaker (Xu, 1997) using a fourth-order polynomial (A1: $\Delta 31.2$ Hz / 0.38 octaves) (Swaminathan, Krishnan, Gandour, & Xu, 2008). Δ represents the difference in f_0 between turning point (65 ms), defined as the minimum f_0 along the duration of the pitch contour, and offset (250 ms) expressed in Hertz and octaves. At the other end of the continuum is a scaled version of the Tone 2 f_0 contour that clearly extends beyond the limits of the normal voice range (A4: $\Delta 299.2$ Hz / 1.99 octaves). Of the two remaining f_0 contours with intermediate rates of acceleration, one represents a pitch pattern that is of marginal relevance linguistically (A2: $\Delta 74.2$ Hz / 0.74 octaves); the other, a pitch pattern that does not occur within a normal voice range (A3: $\Delta 149.2$ Hz / 1.48 octaves). This continuum was derived from a study on the maximum speed at which a speaker can voluntarily change pitch (Xu & Sun, 2002). The maximum velocity in a rising direction was reported to be 61.3 semitones per second (st/s). In this study, we measured velocity from turning point to tonal offset, an excursion time of 185 ms. As compared to 61.3 st/s, A1 (25.4 st/s) falls well within the physiological limits of speed of rising pitch changes; A2 (51.94 st/s) similarly falls under the maximum, but is beginning to approach the maximum limit for changes in rising pitch; both A3 (85.44 st/s) and A4 (129.5 st/s) go far beyond the limit for changes in rising pitch.

The stimuli were generated using click trains whose inter-click intervals were changed according to $1/f_0$ for each polynomial (Supplementary Material: audio files, .mp3; click-

train_stimuli.doc). All stimuli were low pass filtered at 3000 Hz, matched in RMS amplitude, and fixed in duration to 250 ms including a 10 ms rise/fall time (\cos^2 ramps).

The size of the analysis window for each stimulus was fixed at five pitch periods, instead of a constant time interval, resulting in a varying window length demarcated by width of the rectangles (top panel). This process was necessary to ensure the same temporal resolution for the autocorrelation spectral analysis for each stimulus across the pitch acceleration continuum. The center of each window (at 177 ms) corresponded to the *peak* of the pitch acceleration curve (Fig. 3, bottom panel). Note that the pitch acceleration curves differed across stimuli throughout their duration except for the turning point (65 ms).

4.3. Data acquisition

FFR recording protocol and data analysis were similar to those described in previous publications from our laboratory (Bidelman, et al., in press; Krishnan, Gandour, et al., 2009; Krishnan, Swaminathan, et al., 2009). Participants reclined comfortably in an acoustically and electrically shielded booth. They were instructed to relax and refrain from extraneous body movements to minimize movement artifacts. FFRs were recorded from each participant in response to monaural stimulation of the right ear at 80 dB SPL at a repetition rate of 2.76/s. The presentation order of the stimuli was randomized both within and across participants. Control of the experimental protocol was accomplished by a signal generation and data acquisition system (Tucker-Davis Technologies, System III). The stimulus files were routed through a digital to analog module and presented through a magnetically shielded insert earphone (Etymotic, ER-3A).

FFRs were recorded differentially between a non-inverting (positive) electrode placed on the midline of the forehead at the hairline (Fz) and inverting (reference) electrodes placed on (i) the right mastoid (A2); (ii) the left mastoid (A1); and (iii) the 7th cervical vertebra (C7). Another electrode placed on the mid-forehead (Fpz) served as the common ground. FFRs were recorded simultaneously from the three different electrode configurations, and subsequently averaged for each stimulus condition to yield a response with a higher signal-to-noise ratio (Krishnan, Gandour, et al., 2009). All inter-electrode impedances were maintained below 1 k Ω . The EEG inputs were amplified by 200,000 and band-pass filtered from 75 to 1500 Hz (6 dB/octave roll-off, RC, response characteristics). Each response waveform represented the average of 3000 stimulus presentations over a 280 ms analysis window using a sampling rate of 24414 kHz. The experimental protocol took about 100 minutes to complete.

4.4. Data analysis

Short-term autocorrelation functions (ACF) and running autocorrelograms were computed for the FFRs to index variation in FFR periodicities over the duration of the response to time-varying, click-train stimuli. The autocorrelogram (ACG) represents the short term autocorrelation function of windowed frames of a compound signal, i.e., $ACG(\tau, t) = X(t) \times X(t-\tau)$ for each time t and time-lag τ . It is a three dimensional plot quantifying the variations in periodicity and pitch strength as a function of time. The horizontal axis represents the time at which single ACF “slices” are computed while the vertical axis represents their corresponding time lags, i.e., pitch periods. The intensity of each point in the image represents the instantaneous ACF magnitude computed at a given time within the response.

Neural pitch strength was extracted from the FFRs using autocorrelation analysis (Krishnan, Gandour, et al., 2009; Krishnan, Swaminathan, et al., 2009; Swaminathan, Krishnan, Gandour, et al., 2008) applied to a time window centered around the region of maximum

acceleration (Fig. 3, bottom panel). This analysis yielded estimates of both pitch period (time lag associated with the autocorrelation maximum) and pitch strength (magnitude of the normalized autocorrelation expressed as a value between 0 and 1 where, 0 represents an absence of periodicity and 1 represents maximal periodicity). It should be noted here that the time window used for the autocorrelation analysis varied in duration across the stimuli in order to include exactly five pitch periods in the FFR response- a necessary requirement to maintain constant temporal resolution in the analysis process.

Narrow-band spectrograms were obtained from each FFR waveform using a 50 ms Hamming window to evaluate the spectral representation of the harmonics. In order to determine if there were between-group differences in the representation of the pitch relevant harmonics in the FFR, the stimulus spectrum containing the first five harmonics was compared with that of the FFR response spectrum for each stimulus using a spectral correlation analysis confined to the region of maximal acceleration. The resulting correlation coefficient is expressed as a value between -1 and 1. Consistent with the pitch strength measurement, the stimulus time window varied in duration in order to include exactly five pitch periods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research supported by the National Institutes of Health R01 DC008549-01A1 (A.K.) and NIDCD predoctoral traineeship (G.B.). Thanks to Bruce Craig and Duncan Leaf for their assistance with statistical analysis (Purdue Department of Statistics).

References

- Anathanarayan AK, Gerken GM. Post-stimulation effects on the auditory brain stem response partial-masking and enhancement. *Electroencephalography and Clinical Neurophysiology* 1983;55(2):223–226. [PubMed: 6185321]
- Anathanarayan AK, Gerken GM. Response enhancement and reduction of the auditory brain-stem response in a forward-masking paradigm. *Electroencephalography and Clinical Neurophysiology* 1987;66(4):427–439. [PubMed: 2435523]
- Bent T, Bradlow AR, Wright BA. The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance* 2006;32(1):97–103. [PubMed: 16478329]
- Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*. in press.
- Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology* 1996;76(3):1698–1716. [PubMed: 8890286]
- Chandrasekaran B, Gandour JT, Krishnan A. Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. *Restorative Neurology and Neuroscience* 2007;25:195–210. [PubMed: 17942999]
- Gandour JT. Tone perception in Far Eastern languages. *Journal of Phonetics* 1983;11:149–175.
- Howie, JM. *Acoustical studies of Mandarin vowels and tones*. Cambridge University Press; New York: 1976.
- Johnson KL, Nicol TG, Kraus N. Brain stem response to speech: a biological marker of auditory processing. *Ear and Hearing* 2005;26(5):424–434. [PubMed: 16230893]
- Kraus N, Banai K. Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science* 2007;16(2):105–110.

- Kraus N, Nicol T. Brainstem origins for cortical ‘what’ and ‘where’ pathways in the auditory system. *Trends in Neurosciences* 2005;28(4):176–181. [PubMed: 15808351]
- Krishnan, A. Human frequency following response.. In: Burkard, RF.; Don, M.; Eggermont, JJ., editors. *Auditory evoked potentials: Basic principles and clinical application*. Lippincott Williams & Wilkins; Baltimore: 2006. p. 313-335.
- Krishnan A, Gandour JT. The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language* 2009;110:135–148. [PubMed: 19366639]
- Krishnan A, Gandour JT, Bidelman GM. Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research* 2010a;1313:124–133. [PubMed: 19961835]
- Krishnan A, Gandour JT, Bidelman GM. The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics* 2010b;23:81–95. [PubMed: 20161561]
- Krishnan A, Gandour JT, Bidelman GM, Swaminathan J. Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport* 2009;20(4):408–413. [PubMed: 19223793]
- Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience* 2009;21(6):1092–1105. [PubMed: 18702588]
- Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research. Cognitive Brain Research* 2005;25(1):161–168. [PubMed: 15935624]
- Langner G. Periodicity coding in the auditory system. *Hearing Research* 1992;60(2):115–142. [PubMed: 1639723]
- Langner, G. Topographic representation of periodicity information: The 2nd neural axis of the auditory system.. In: Syka, J.; Merzenich, M., editors. *Plasticity of the central auditory system and processing of complex acoustic signals*. Plenum Press; New York: 2004. p. 21-26.
- Luo H, Boemio A, Gordon M, Poeppel D. The perception of FM sweeps by Chinese and English listeners. *Hearing Research* 2007;224(1-2):75–83. [PubMed: 17207949]
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9:97–113. [PubMed: 5146491]
- Patel AD, Iversen JR. The linguistic benefits of musical abilities. *Trends in Cognitive Sciences* 2007;11(9):369–372. [PubMed: 17698406]
- Schwartz DA, Purves D. Pitch is determined by naturally occurring periodic sounds. *Hearing Research* 2004;194(1-2):31–46. [PubMed: 15276674]
- Swaminathan J, Krishnan A, Gandour JT. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport* 2008;19(11):1163–1167. [PubMed: 18596621]
- Swaminathan J, Krishnan A, Gandour JT, Xu Y. Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Transactions on Biomedical Engineering* 2008;55(1):281–287. [PubMed: 18232372]
- Tzounopoulos T, Kraus N. Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 2009;62(4):463–469. [PubMed: 19477149]
- Wong PC, Perrachione TK. Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics* 2007;28(4):565–585.
- Xiong Y, Zhang Y, Yan J. The neurobiology of sound-specific auditory plasticity: A core neural circuit. *Neuroscience and Biobehavioral Reviews* 2009;33(8):1178–1184. [PubMed: 19014967]
- Xu Y. Contextual tonal variations in Mandarin. *Journal of Phonetics* 1997;25:61–83.
- Xu Y, Krishnan A, Gandour JT. Specificity of experience-dependent pitch representation in the brainstem. *Neuroreport* 2006;17(15):1601–1605. [PubMed: 17001276]
- Xu Y, Sun X. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 2002;111(3):1399–1413. [PubMed: 11931317]
- Yip, M. *Tone*. Cambridge University Press; New York: 2003.
- Zatorre RJ, Gandour JT. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 2008;363(1493):1087–1104.

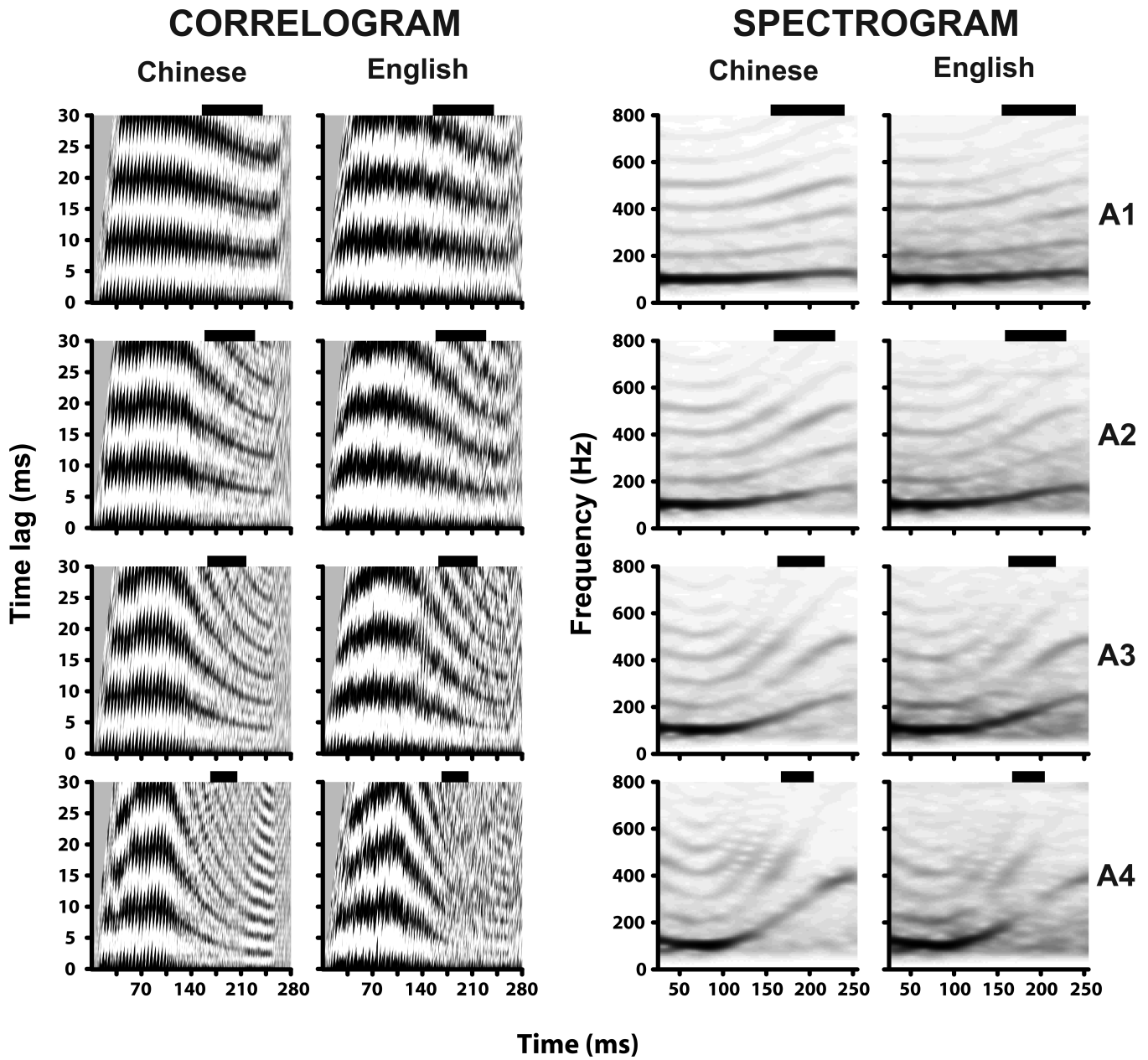


Figure 1. Average correlograms (columns 1-2) and spectrograms (columns 3-4) derived from the full length (280 ms) FFR waveforms of Chinese and English groups in response to Mandarin Tone 2 (A1), and three modifications of Tone 2 with increasing rates of pitch acceleration (A2, A3, A4). At all acceleration levels, correlograms of the Chinese group show clearer, tighter bands of temporal regularity (black) in FFR phase-locked activity at the fundamental period ($1/f_0$) and its multiples as compared to the English group. This is especially noteworthy in the temporal analysis window of maximum pitch acceleration centered at 177 ms. Black horizontal bars indicate the span of the analysis window per stimulus (see also Fig. 3). Similarly, at all acceleration levels, the spectrograms for the Chinese group show more robust, clearer harmonics than the English group.

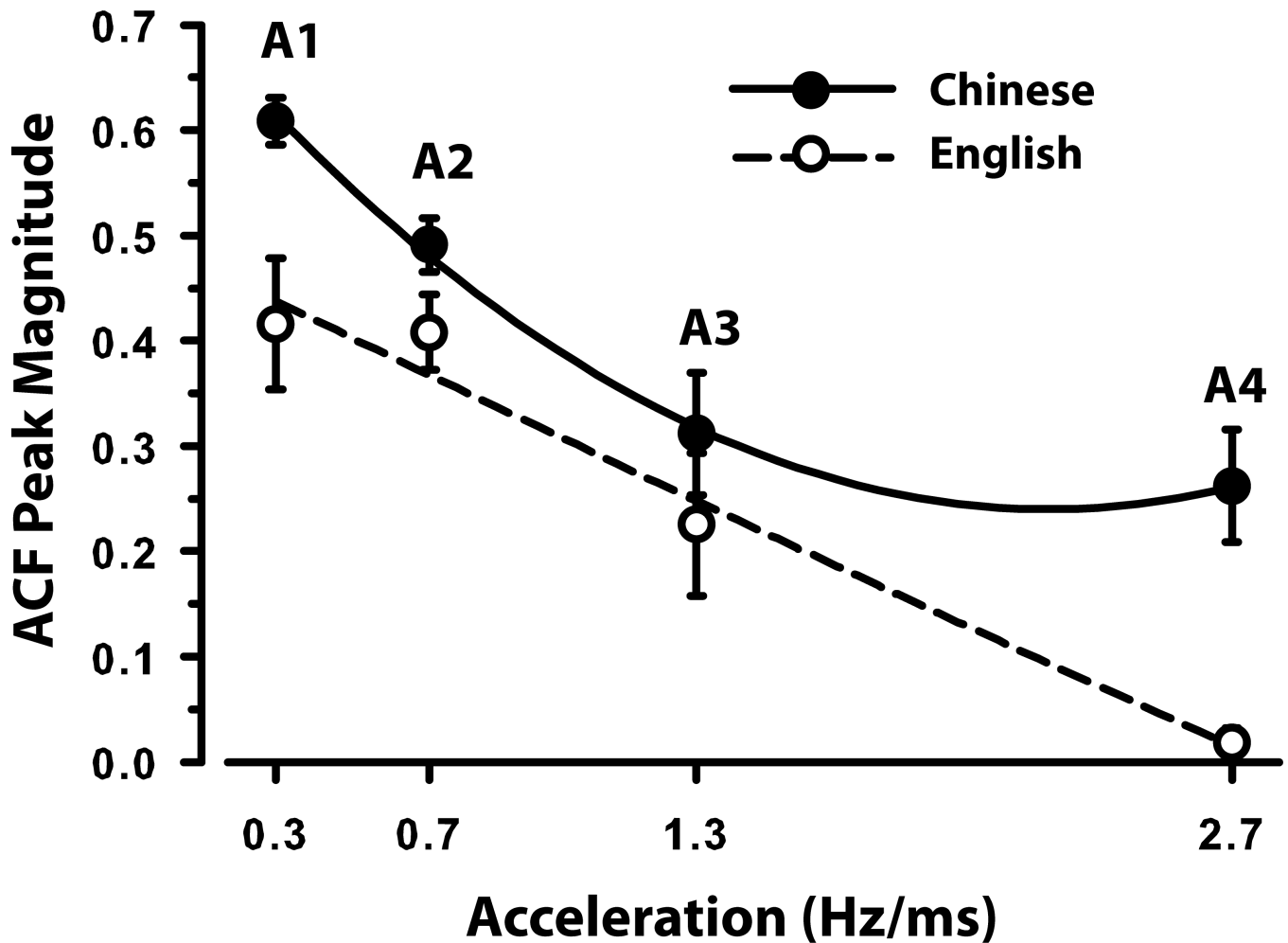


Figure 2.

Group comparisons of pitch strength within the region of maximum acceleration of the stimulus derived from FFR responses to click-train stimuli as a function of pitch acceleration. FFR pitch strength, as measured by the magnitude of the normalized autocorrelation peak, of the Chinese group is greater than that of the English in response to Mandarin Tone 2 (A1, 0.3 Hz/ms) as well as to a pitch pattern that does not occur in natural speech (A4, 2.7 Hz/ms). In the English group, pitch strength shows a steady, steep decline across the continuum, approaching zero, i.e., the absence of a phase-locked response, at its opposite end. In the Chinese group, on the other hand, pitch strength exhibits a more gradual decline, but never approaches zero. Instead, pitch strength begins to level off once a pitch pattern moves clearly beyond the normal voice range (A3, 1.3 Hz/ms). Solid lines show quadratic fits to the data; error bars = ± 1 SE.

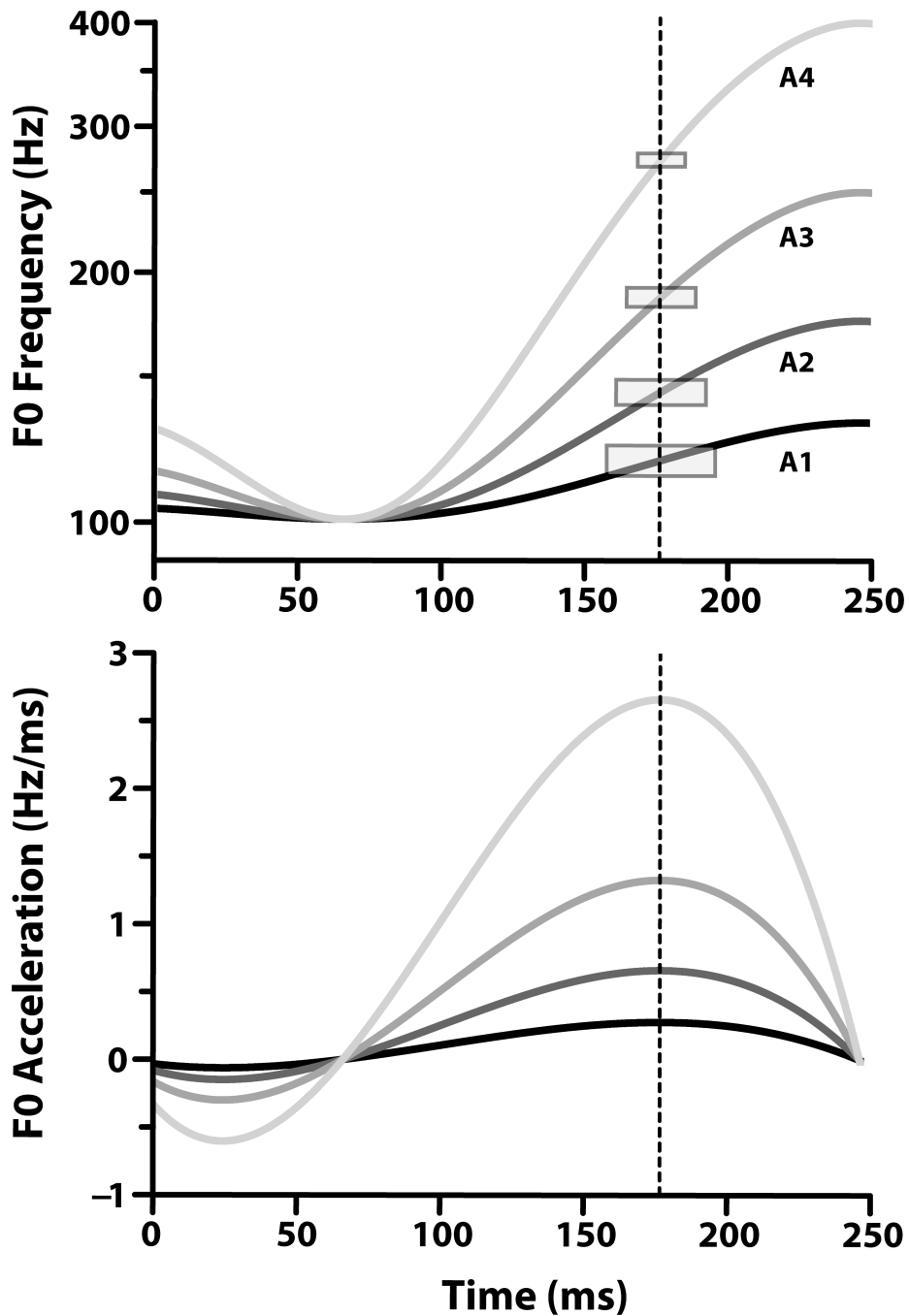


Figure 3.

Time-varying click-train stimuli used to evoke brainstem responses to f_0 patterns that are differentiated by varying degrees of rising acceleration. f_0 contours of all four stimuli are displayed on a logarithmic scale spanning two octaves - from 100.8 Hz, the minimum stimulus frequency, to 400 Hz (top panel). These stimuli represent a continuum of rates of acceleration from an exemplary Tone 2 rate in natural speech to an f_0 rate that falls well beyond the normal voice range. One end of the continuum is represented by the f_0 contour of Mandarin Tone 2 as produced in citation form (A1). At the opposite end is an f_0 contour that clearly falls outside the bounds of the normal voice range (A4). The remaining two f_0 contours that deviate from the prototypical Tone 2 reflect intermediate rates of acceleration.

One represents a pitch pattern that is of marginal relevance linguistically (A2); the other a pitch pattern that does not occur within a normal voice range (A3). The vertical dotted line at 177 ms defines the center of the analysis window for each stimulus. The location of this line is calculated from the maximum pitch acceleration (bottom panel). A1, A2, A3, A4 = rates of pitch acceleration.

Table 1

Stimulus-to-response spectral correlation coefficients per language group as a function of stimulus acceleration rate

Stimulus	Group	
	Chinese	English
A1	0.60 (0.04)	0.35 (0.05)
A2	0.54 (0.04)	0.33 (0.05)
A3	0.43 (0.05)	0.32 (0.06)
A4	0.46 (0.06)	0.19 (0.04)

Note. Values are expressed as mean and standard error (in parentheses).