Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Brain Research 25 (2005) 161 – 168
COGNITIVE BRAIN RESEARCH
www.elsevier.com/locate/cogbrainres

Research Report

Encoding of pitch in the human brainstem is sensitive to language experience

Ananthanarayan Krishnan[a], Yisheng Xu[a], Jackson Gandour[a,*], Peter Cariani[b]

[a]Department of Speech, Language, and Hearing Sciences, Purdue University, Heavilon Hall 1353, 500 Oval Drive, West Lafayette, IN 47907-2038, USA
[b]Eaton Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

Accepted 3 May 2005
Available online 1 June 2005

Abstract

Neural processes underlying pitch perception at the level of the cerebral cortex are influenced by language experience. We investigated whether early, pre-attentive stages of pitch processing at the level of the human brainstem may also be influenced by language experience. The human frequency following response (FFR), reflecting sustained phase-locked activity in a population of neural elements, was used to measure activity within the rostral brainstem. FFRs elicited by four Mandarin tones were recorded from native speakers of Mandarin Chinese and English. Pitch strength (reflecting robustness of neural phase-locking at the pitch periods) and accuracy of pitch tracking were extracted from the FFRs using autocorrelation algorithms. These measures revealed that the Chinese group exhibits stronger pitch representation and smoother pitch tracking than the English group. Consistent with the pitch data, FFR spectral data showed that the Chinese group exhibits stronger representation of the second harmonic relative to the English group across all four tones. These results cannot be explained by a temporal pitch encoding scheme which simply extracts the dominant interspike interval. Rather, these results support the possibility of neural plasticity at the brainstem level that is induced by language experience that may be enhancing or priming linguistically relevant features of the speech input.

Theme: Sensory systems
Topic: Auditory systems: central physiology

Keywords: Auditory; Brainstem; Speech; Frequency following response; Experience-dependent plasticity; Mandarin Chinese tone

1. Introduction

Languages that exploit variations in pitch to signal meaning differences in monosyllabic words (e.g., Mandarin Chinese: *ma*$^{\text{high level}}$ 'mother', *ma*$^{\text{high rising}}$ 'hemp', *ma*$^{\text{low falling rising}}$, 'horse', *ma*$^{\text{high falling}}$ 'scold') are called *tone languages*. Language processing is known to be lateralized to the left hemisphere, whereas pitch perception is mediated in the right hemisphere [41]. In tone perception, cross-language behavioral [38], neuropsychological [9], and neuroimaging [10,17] studies reveal a leftward asymmetry for native speakers of tone languages. At the cortical level, these data clearly suggest that the neural substrates of pitch perception in the processing of lexical tones are shaped by language experience. Moreover, it has also been shown that language experience may even influence basic auditory processes (e.g., pure tone perception) at the level of auditory cortex [30,37].

This experience-dependent neural plasticity is not limited to the auditory cortex. Suga and his co-workers have demonstrated the changes in the response properties and frequency maps in the inferior colliculus of bats following auditory conditioning or focal electrical stimulation of the auditory cortex [34–36,40]. Auditory experience of altered interaural cues for localization in young owls has been shown to produce frequency-dependent changes in interaural time difference tuning and frequency tuning of neurons in the inferior colliculus [12,19]. In humans, a shortening of wave V latency, presumably generated in the inferior colliculus of the auditory brainstem, has been reported in a group of hearing-impaired listeners following

* Corresponding author. Fax: +1 765 494 0771.
E-mail address: gandour@purdue.edu (J. Gandour).

0926-6410/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2005.05.004

the use of amplification as compared to no changes in wave V latency for a control group of hearing-impaired listeners who did not use amplification [27]. More directly relevant to this study are the improvements reported in encoding of the human frequency following response (FFR), the IC also being its presumed generator site, following auditory training of children with learning impairment [29]. As far as we know, it has yet to be demonstrated that neural plasticity in the FFR can be attributed to language experience.

While it is important to identify language-dependent processing systems at the cortical level, a complete understanding of the neural organization of language can only be achieved by viewing language processes as a set of computations or mappings between representations at different stages of processing [15]. In speech perception, for example, early processing stages are not to be dismissed as auditory areas and not relevant to language processing. Rather, early stages of processing on the input side may perform computations on the acoustic data that are relevant to *linguistic* as well as non-linguistic auditory perception. The degree of linguistic specificity is yet to be determined for computations performed at the level of the auditory brainstem.

The fact that the primary acoustic correlate of lexical tone is voice fundamental frequency (F_0) [8] provides a window for exploring processing of the same acoustic parameter at two different stages of the language processing system. It is well-known that discharge periodicities and interspike intervals related to F_0 are present in the responses of auditory nerve fibers [25,26]. Neural phase-locking related to F_0 plays a dominant role in the encoding of low pitch associated with complex sounds [4]. The scalp-recorded human FFR reflects sustained phase-locked activity in a population of neural elements within the rostral brainstem [11,24,32]. It has been demonstrated that the human FFR preserves certain spectrum-relevant information of speech sounds [6,7,20–22,28], and moreover, pitch-relevant information about complex sounds that yield *time-invariant* pitch [13]. This pitch-relevant neural activity appears to be based on the temporal pattern of neural activity in the brainstem and not simply a reflection of neural synchronization to waveform envelope modulation pattern [14]. Indeed, the human FFR has been shown to be sufficiently dynamic to encode *time-varying* pitch of the four lexical tones of Mandarin Chinese [23].

The aim of this cross-language FFR study is to determine whether pitch encoding at the brainstem level is language-dependent in its response properties. FFRs are elicited in response to the four Mandarin tones. By comparing native speakers of a tone language (Mandarin) to those of a non-tone language (English), we are able to determine the extent to which these response properties (pitch strength, tracking accuracy) are sensitive to language experience.

2. Materials and methods

2.1. Subjects

Fourteen adult native speakers of Mandarin and 13 native speakers of American English, ranging in age from 21 to 27 years, participated in the study. All Chinese subjects were born and raised in mainland China and classified as late Mandarin–English bilinguals, not having received formal instruction in English until the age of 11. They all resided in the USA for at least 1 but not more than 4 years. Hearing sensitivity in all subjects was better than 15 dB HL for octave frequencies from 500 to 8000 Hz. All subjects gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

2.2. Stimuli

A set of Mandarin monosyllables was chosen to contrast the four lexical tones: /*yi*^{1}/ 'clothing', /*yi*^{2}/ 'aunt', /*yi*^{3}/ 'chair', /*yi*^{4}/ 'easy'. This quadruplet is especially advantageous for isolating effects of encoding of voice pitch at the level of the auditory brain stem. The four words are minimally distinguished by tone (*/1, 2, 3, 4/*); onset (*/y/*) and rhyme (*i*) are identical. All four tones exhibit voice fundamental frequency (F_0) trajectories and harmonics that lie within the range of easily recordable FFRs (below about 2000 Hz). This tonal inventory includes F_0 trajectories that exhibit changes in direction of pitch movement (unidirectional rising = Tone 2; unidirectional falling = Tone 4; bidirectional falling–rising = Tone 3) as well as relatively steady pitch movement (level = Tone 1).

Syllables were created using a synthesis-by-rule scheme [18]. The fixed parameters were duration, voice amplitude, and vowel formant frequencies. Duration was fixed at 250 ms. Amplitude was constant at 60 dB. Vowel formant frequencies were steady state (in Hz): $F_1 = 300$; $F_2 = 2500$; $F_3 = 3500$; and $F_4 = 4530$ [16]. F_0 contours of the four Mandarin tones were modeled after data from [39]. Thus, we were able to evaluate FFRs elicited by different F_0 contours in a fixed spectral context (Fig. 1).

2.3. Experimental protocol

Subjects reclined comfortably in an acoustically and electrically shielded booth. They were instructed to relax and refrain from extraneous body movements to minimize movement artifacts. All stimuli were presented monaurally at 60 dB nHL at a repetition rate of 3.13/s The order of stimuli (*yi*^{1}, *yi*^{2}, *yi*^{3}, *yi*^{4}) was randomized across subjects. All stimuli were controlled by a signal generation and data acquisition system (Tucker-Davis Technologies, System II). The stimulus files were routed through a digital to analog module and presented monaurally to each ear through magnetically shielded insert earphones (Biologic, TIP-300).

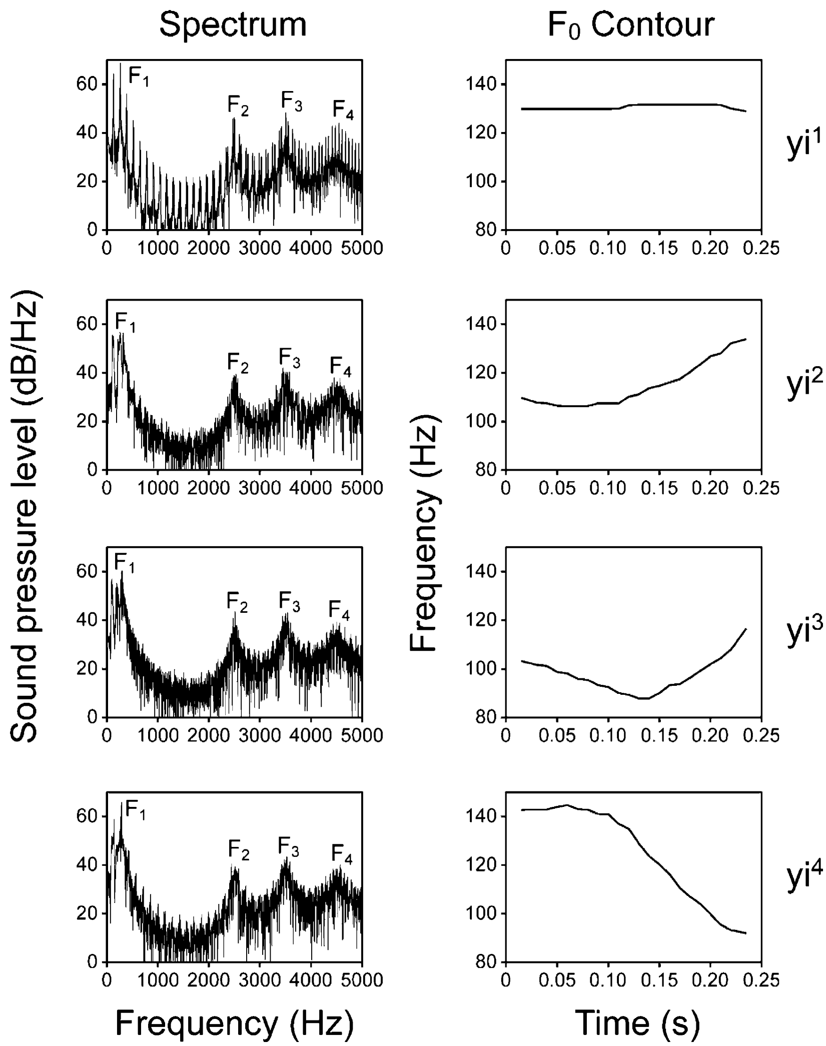

Fig. 1. Acoustic spectra (left column) and F_0 contours (right column) of the synthetic speech stimuli. Each spectrum contains four steady-state formants (F_1, F_2, F_3, F_4). Spectra are identical across all four stimuli. F_0 contours vary depending on the Mandarin tonal category (1, 2, 3, 4).

For all four test conditions, FFRs were recorded from each subject to monaural stimulation of the right ear and the left ear. These evoked responses were recorded differentially between scalp electrodes placed on the midline of the forehead at the hairline and the 7th cervical vertebra (C7). Another electrode placed on the mid-forehead (Fpz) served as the common ground. The inter-electrode impedances were maintained below 3000 Ω. The EEG inputs were amplified by 200,000 and band-pass filtered from 100 to 3000 Hz (6 dB/octave roll-off, RC response characteristics). Each response waveform represents an average of 2000 stimulus presentations over a 260-ms analysis window using a sampling rate of 20 kHz. FFR data from each subject were recorded in two 90 min sessions.

2.4. Analysis procedure

The ability of the FFR to follow the pitch change in the stimuli was evaluated by extracting the F_0 contour from the grand-average FFRs using a periodicity detection short-term autocorrelation algorithm [3]. Essentially, this algorithm performs a short-term autocorrelation analysis on a number of small segments (frames) taken from the signal (stimuli and FFR). This analysis yielded estimates of both pitch period (time lag associated with the autocorrelation maximum) and pitch strength (magnitude of the normalized autocorrelation peak expressed as harmonic-to-noise ratio ranging from 0 to 1).

Average autocorrelation magnitude (pitch strength) was derived from all the time frames in each FFR waveform. A mixed model ANOVA, with subject as a random effect nested within group (Chinese, English) and stimulus and ear as repeated measures, was conducted on the average autocorrelation magnitude to determine whether pitch strength varied as a function of language experience in relation to tonal stimulus and stimulated ear. A similar measure of pitch salience has demonstrated good correspondence between the normalized magnitude of the autocorrelation peak and perceived pitch salience for a number of low pitch complex sounds [4].

Short-term autocorrelation functions and running autocorrelograms were computed for the stimuli and the grand-average FFRs to index variation in FFR periodicities over the duration of the response. The autocorrelogram is an expansion of the signal that plots post-stimulus onset time vs. time lag, i.e., $ACG(\tau,t) = X(t) \times X(t-\tau)$ for each time t and time lag τ. Thus, it represents the running distribution of all-order intervals present in the population response [4,23].

Crosscorrelation analysis was performed between the F_0 contours extracted from the original speech stimuli and those derived from the FFR waveforms to index the degree to which the FFR followed the stimulus time structure. Crosscorrelation coefficients were rank transformed in order to improve the normality of the data. A mixed model ANOVA, with subject as a random effect nested within group and stimulus and ear as repeated measures, was conducted on the rank of F_0 crosscorrelation to evaluate the effects of language experience on the ability of the FFR to follow time-varying F_0 information in the Mandarin tonal stimuli.

Narrow-band spectrograms were obtained from each FFR waveform using a 30-ms analysis window (Gaussian) to evaluate the spectral composition and magnitude of the phase-locked neural activity at each of the first five harmonics. Twenty short-term spectral slices were derived from the FFR spectrogram at every 10 ms between 30 ms and 220 ms. H_{1-5} harmonic peaks were identified for each spectral slice at positions close to the corresponding harmonics of the stimulus. Magnitude of each harmonic was expressed in dB relative to each harmonic's noise floor. Thus, the mean harmonic magnitude represents the harmonic-to-noise ratio (HNR) in dB computed from the twenty slices. These data were subjected to a mixed model ANOVA (subjects as a random effect) for the effects of group, ear, tonal stimulus, and harmonic to determine whether spectral representation of individual harmonics varies depending on language experience.

3. Results

3.1. Representation of voice pitch

Short-term autocorrelation functions and the running autocorrelograms of the FFR to the Tone 2 stimulus (*yi*^{2}) are shown in Fig. 2 for the Chinese and English groups. In the autocorrelation functions (left panels), a peak at the fundamental period $1/F_0$ is observed for both groups, which means that phase-locked activity to the fundamental period is present regardless of language experience. However, the peak for the English group is smaller and broader relative to the Chinese group, suggesting that phase-locked activity is not as robust for English listeners. In the autocorrelograms (right panels), a time-variant band of phase-locked activity (white) closely follows the decreasing fundamental period, corresponding to increasing F_0 in *yi*^{2}, especially over the last half of its duration (cf. Fig. 1). Consistent with their respective autocorrelation functions, the band of phase-locked interval for the Chinese group is narrower than that for the English, suggesting that phase-locked activity for the

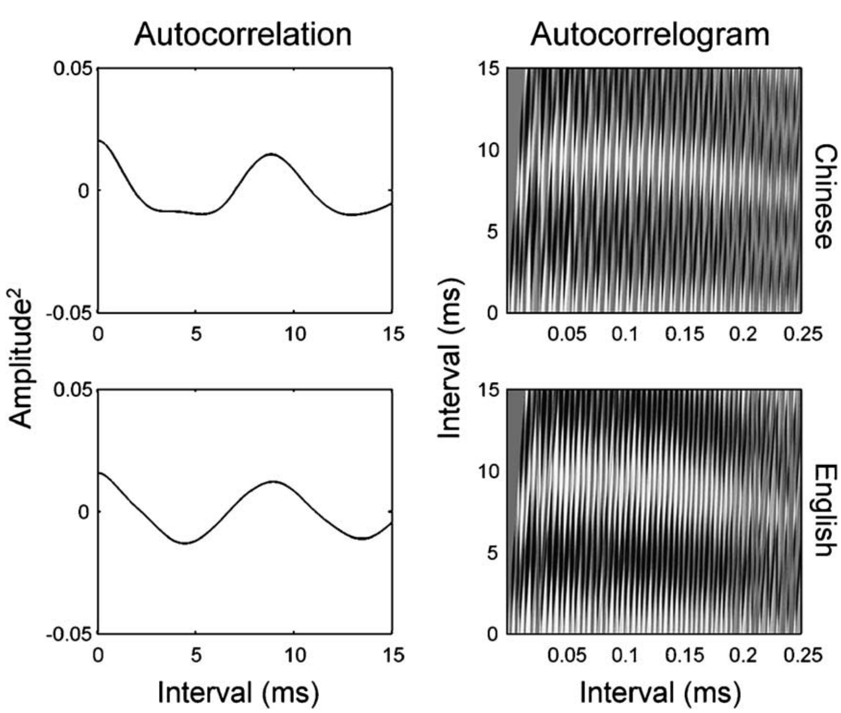

Fig. 2. Short-term autocorrelation functions (left panels) and running autocorrelogram (right panels) of the average FFR waveforms of Chinese (top panels) and English (bottom panels) groups when listening to the Tone 2 stimulus (*yi*^{2}). The broader phase-locked pitch interval band (white) for the English group is consistent with the broader and smaller magnitude autocorrelation peak.

Chinese listeners is, not only more robust, but also more accurate than that of English listeners.

FFR pitch strength, as measured by average autocorrelation magnitude, is shown for each of the four Mandarin tones by language group in Fig. 3. Pitch strength was significantly greater for the Chinese group than for the English ($F_{1,21}$ = 10.05, P = 0.0041). There was no group × stimulus interaction, meaning that pitch strength was higher for the Chinese group than the English across all four Mandarin tones. The only significant effect involving ear difference was the ear × stimulus interaction ($F_{3,45}$ = 3.86, P = 0.0153). Tests of simple main effects show that the Tone 3 stimulus (yi^3) elicits a significant right ear advantage for both Chinese and English groups ($F_{1,45}$ = 6.72, P = 0.0128).

Grand-average F_0 contours extracted from FFRs are displayed for the Tone 2 stimulus (yi^2) in both groups in Fig. 4. No matter Chinese or English, the FFR response is observed to generally follow the rising F_0 contour of the original speech stimulus. However, pitch tracking is more variable (see enlarged inset) for the English group compared to the Chinese group. Consistent with this observation, FFR pitch tracking, as measured by rank-transformed cross-correlation coefficients, was significantly greater (i.e., more accurate) in the Chinese group across all four Mandarin tones (Fig. 5) as compared to the English group ($F_{1,24}$ = 17.92, P = 0.0003).

3.2. Representation of spectral components

Narrow-band spectrograms of the original speech stimulus (left panel) and of the grand-average FFRs in response to the Tone 2 (yi^2) stimulus for the Chinese (middle panel) and English (right panel) groups are displayed in Fig. 6. As expected, the yi^2 stimulus spectrogram reveals energy bands at several multiples of F_0 including stronger energy bands at F1-related harmonics (h_2, h_3). The FFR spectrograms show energy bands at several harmonics for both groups. The energy band of the second harmonic (h_2), in particular, appears to be stronger in the Chinese group relative to the English group with little or no difference in the energy bands at the higher harmonics.

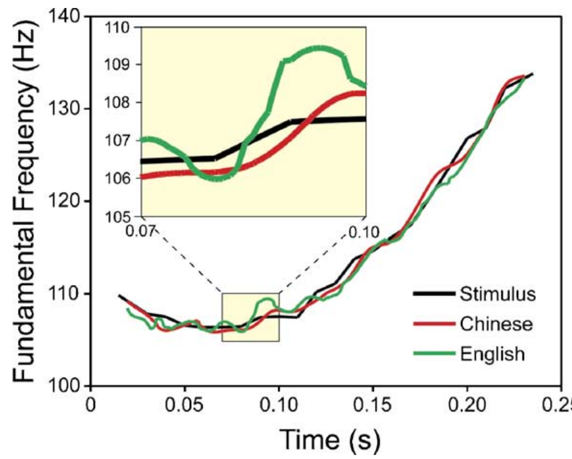

Fig. 4. Grand-average F_0 contours of Tone 2 (yi^2) derived from the FFR waveforms of all subjects across both ears in the Chinese (red) and English (blue) groups. The F_0 contour of the original speech stimulus (yi^2) is displayed in black. The enlarged inset shows that the F_0 contour derived from the FFR waveforms of the Chinese group more closely approximates that of the original stimulus (yi^2) when compared to the English group.

Results of an ANOVA (group × harmonic × stimulus × ear) on harmonic magnitude revealed a significant group × harmonic interaction ($F_{4,96}$ = 6.35; P = 0.0001). No higher-order three-way interactions were significant. As shown in Fig. 7, a significantly higher harmonic magnitude was observed for the Chinese listeners compared to English across all four Mandarin tones for h_2 only ($F_{1,96}$ = 27.99; p < 0.0001).

4. Discussion

4.1. Representation of voice pitch

The major finding of this study is that greater pitch strength and more accurate pitch tracking of linguistically relevant pitch contours occur at the level of the auditory brainstem for native listeners of a tone language as

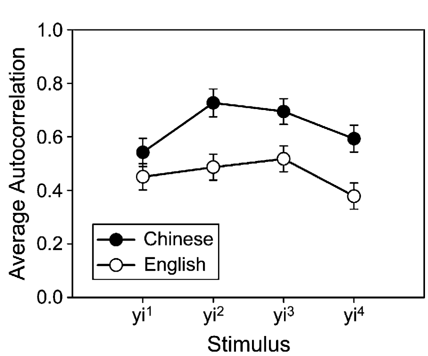

Fig. 3. Comparison between Chinese and English groups on the autocorrelation magnitude for each of the four speech stimuli. Higher magnitudes are observed in the Chinese group relative to the English regardless of tone.

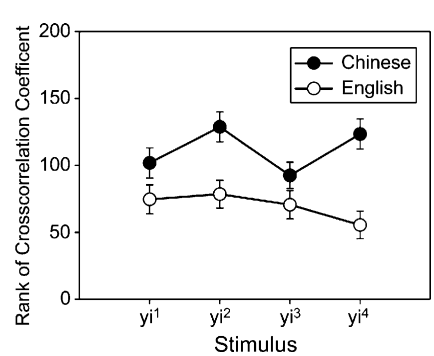

Fig. 5. Comparison between Chinese and English groups on the rank-transformed crosscorrelation coefficient between the F_0 contours of the four tonal stimuli and FFR waveforms. Higher crosscorrelation coefficients are observed in the Chinese group relative to the English regardless of tone.

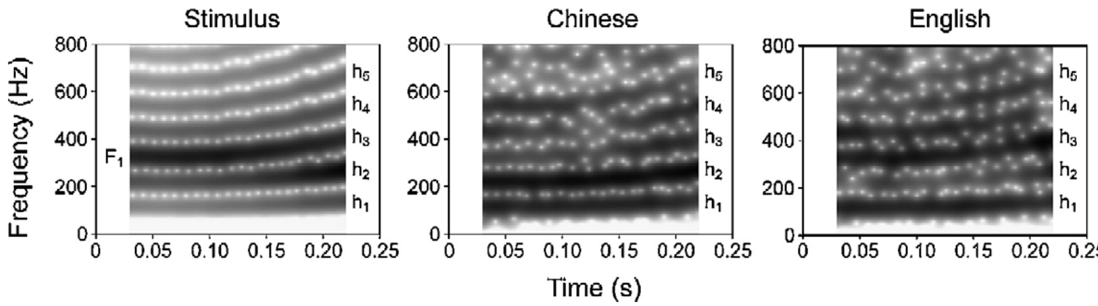

Fig. 6. Narrow-band spectrograms of the original Tone 2 stimulus (*yi*[2]) (left panel) and of the grand-average FFR waveforms of the Chinese (middle panel) and English (right panel) groups. F_1 identifies the first formant; h_1-h_5 the first through fifth harmonics.

compared to non-native listeners. In terms of the temporal pattern of neural activity, this means that the degree of phase-locking is greater and the variability is smaller around the phase-locked interval for the Chinese listeners compared to the English listeners. Current temporal encoding schemes of pitch extraction, based on the dominant interval in the distribution of interspike intervals, rely purely on the acoustic properties of the stimulus. Consequently, they would predict no significant differences in the characteristics of encoding (i.e., pitch strength, accuracy of pitch tracking) across listeners regardless of language experience. The present findings, however, clearly demonstrate that the encoding scheme is not static nor is it dedicated to faithfully extract only the physical properties of the stimulus. Rather, they are consistent with a temporal encoding scheme which is plastic, i.e., sensitive to language experience. This plasticity enables enhancing or priming of temporal intervals that carry linguistically relevant features of pitch contours. The relatively greater pitch strength and smoother pitch tracking in native Mandarin listeners may reflect the operation of this language-dependent encoding scheme.

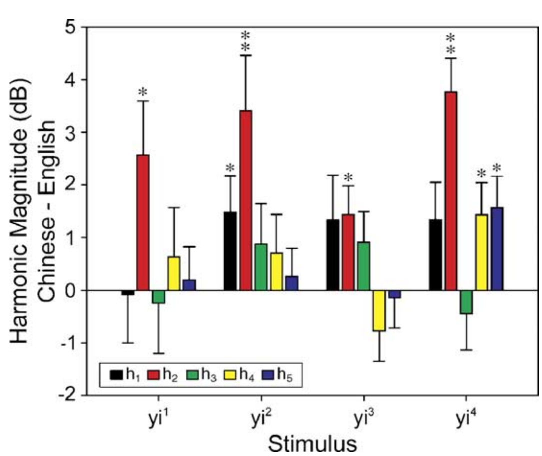

Fig. 7. Cross-language differences in FFR harmonic magnitude between the two groups (Chinese, English) for the first five harmonics (h_1-h_5) in response to each of the four Mandarin tones. Positive and negative bars indicate greater harmonic magnitude for the Chinese and English groups, respectively. Pooling across both ears, the Chinese group show significantly greater harmonic magnitude for h_2 (red positive bars) regardless of tonal stimulus. Error bars represent standard errors. $*P < 0.05$; $**P < 0.01$.

One possible encoding scheme is that the phase-locking at the pitch period is enhanced by both an active facilitation/disinhibition of the pitch-relevant intervals and/or inhibition of phase-locking at intervals other than the pitch period. At several levels in the human auditory system, excitatory and inhibitory neural interactions are known to play an important role in enhancing signal selection [1,2]. Also relevant are the different types of subcortical neurons that are sensitive to specific values of the parameters that characterize species-specific animal sounds [33,34]. These neurons are subject to corticofugal egocentric selection along several stimulus parameters. Whether the language-dependent effects observed for the Chinese listeners reflect local processes in the brainstem and/or corticofugal modulation of neural responses [35,36] cannot be determined from the current experimental design. Nevertheless, it is well-established that this corticofugal system is influenced by auditory experience [36]. Given that corticofugal egocentric selection is short term and takes time (latency) to be activated and that the FFR-related changes occur within 6–8 ms, we speculate that the operation of this language-dependent encoding scheme, presumably induced by experience with Mandarin tones, is local to the generator of the FFR in the auditory brainstem. These local brainstem mechanisms could mediate experience-dependent plasticity by altering the balance between excitation and inhibition in inferior colliculus neurons. Such a possibility is foreshadowed by earlier findings from the animal literature, i.e., changes in the spatial map in the inferior colliculus of bats following alteration of interaural cues [12,19] and persistence of inferior colliculus plasticity in mice after deactivation of the corticofugal system [40].

As a factor additional to the presence of tonal cues, we acknowledge that a word familiarity bias is a potential confound due to long-term learning experience. We argue, however, that it is unlikely that this factor can account for our experimental outcomes. Assuming that word familiarity is the driving force, one would expect to see more robust harmonic representations across the board (h_1-h_5) instead of enhanced representation of the second harmonic only. This finding is consistent with the view that the relatively lower harmonics, in this case, h_2, play a prominent role in the representation of pitch information in the FFR. Furthermore, in our first FFR study of tonal processing [23], we presented two stimuli that were spectrally different but had identical F_0 contours: *yi*[3] (speech) and *hum*[3] (nonspeech). Interestingly, the FFR autocorrelograms for these two stimuli are virtually identical, indicating that they both produced equivalent pitch percepts in spite of the fact that *yi*[3] is a familiar word, whereas *hum*[3] is not recognizable as a Chinese word.

Using steady-state complex sounds, it has been demonstrated that the phase-locked activity corresponding to F_0 is not simply the result of neural synchronization to the stimulus envelope modulation pattern [14] but rather represented pitch-relevant temporal discharge patterns of neurons in the upper auditory brainstem pathway [13]. These earlier observations as well as the FFR data presented herein lend further support to the dominant interval hypothesis of pitch encoding. The presence of a prominent interval band in the phase-locked FFR neural activity that closely follows the fundamental period ($1/F_0$) strongly supports a robust temporal code for voice pitch. This prominent interval is based on temporal distribution of phase-locked neural activity in a population of neural elements in the rostral brainstem [13]. Moreover, our FFR data are consistent with physiologic studies that show that the most frequent interspike interval, i.e., the dominant interval in the auditory nerve population interval distribution, corresponds closely to the fundamental period and the low pitch of a variety of complex stimuli [4]. The strength of this prominent interval may vary in response to voice pitch depending on whether the pitch patterns carry linguistically relevant information specific to information-bearing units of a particular language. In this study, pitch patterns at the syllable level are linguistically significant in the Chinese language only. We are not saying that the dominant interval for pitch shifts from language to language but rather that its weighting may vary depending on a listener's implicit knowledge of pitch variations associated with specific linguistic functions (e.g., lexical tone).

4.2. Representation of spectral components

We further observe that the temporal distribution of the phase-locked activity to individual harmonics differs as a function of language experience. Our results show, not only a clear dominance of the second harmonic for all stimuli, but also a trend toward better representation of multiple harmonics in the Chinese group compared to the English. This finding complements our data on voice pitch representation. That is, the stronger pitch and more accurate pitch tracking in the Chinese group co-occur with a relatively stronger representation of pitch-relevant harmonics. The weaker representation of higher harmonics may indicate that they are not as important to the salience of linguistically relevant pitch patterns associated with Mandarin tones. This interpretation is consistent with the fact that the English group shows weaker pitch and less accurate pitch tracking co-occurring with relatively weaker representation of pitch-relevant harmonics across tones. Moreover, psychoacoustic and physiologic studies indicate that complex stimuli produce stronger and more accurate pitch percepts when lower harmonics in the pitch dominant region are presented [4,5,31]. Finally, cross-language differences point to a pitch perception bias that may be enhanced when listeners are presented with behaviorally relevant auditory signals.

4.3. Directions for future research

The long-term goal of this research program is to characterize the nature of neural processing of voice pitch in various aspects of speech prosody at the level of the auditory brainstem. In our first experiment, we demonstrated that the human FFR is sufficiently dynamic to encode time-varying pitch of the four lexical tones of Mandarin Chinese [23]. In this second experiment, we demonstrate for the first time that neural activity in the rostral brainstem is sensitive to language experience, i.e., language-dependent. At this point, a question arises as to whether these observed FFR effects are speech-specific or do they extend to all periodic stimuli including nonspeech sounds. The third experiment already underway in our laboratory addresses this question by comparing FFR responses of Chinese and English listeners to tonal contours representative of all four Mandarin tones and homologous nonspeech F_0 contours. Another question that arises is whether these observed FFR effects are language-specific or do they extend to other tone languages. This question is to be addressed in a fourth experiment by comparing FFR responses of Chinese and Thai listeners to Mandarin and Thai tonal contours. Other aspects of speech prosody, linguistic and affective, are to be investigated in non-tone as well as tone languages. Taken together, these findings promise to offer fresh insights into the neurobiology of pitch perception at the level of the auditory brainstem.

5. Conclusions

The scalp-recorded FFR provides a non-invasive window to view neural processing of voice pitch in human speech sounds at the level of the auditory brainstem. Our findings demonstrate that experience-driven adaptive neural mechanisms are involved subcortically that sharpen response properties of neurons tuned for processing pitch contours that are of special relevance to a particular language. From the perspective of auditory neuroethology, this adjustment in processing pitch contours of Mandarin tones is comparable to neural mechanisms that are developed for processing behaviorally relevant sounds (e.g., communication, orientation) in other non-primate and non-human primate animals. Auditory processing at this level may not be limited to a simple representation of acoustic features of speech stimuli. Indeed, language-dependent or even speech-specific operations may begin before the signal reaches the cerebral cortex.

Acknowledgments

This research was supported in part by research grants to JG from the National Institutes of Health (R01 DC04584-04) and the Purdue Research Foundation.

References

[1] A.K. Ananthanarayan, G.M. Gerken, Post-stimulation effects on the auditory brain stem response partial-masking and enhancement, Electroencephalogr. Clin. Neurophysiol. 55 (1983) 223–226.

[2] A.K. Ananthanarayan, G.M. Gerken, Response enhancement and reduction of the auditory brain-stem response in a forward-masking paradigm, Electroencephalogr. Clin. Neurophysiol. 66 (1987) 427–439.

[3] P. Boersma, Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound., Proc. Inst. Phon. Sci. 17 (1993) 97–110.

[4] P.A. Cariani, B. Delgutte, Neural correlates of the pitch of complex tones: I. Pitch and pitch salience, J. Neurophysiol. 76 (1996) 1698–1716.

[5] H. Fastl, G. Stoll, Scaling of pitch strength, Hear. Res. 1 (1979) 293–301.

[6] G.C. Galbraith, P.W. Arbagey, R. Branski, N. Comerci, P.M. Rector, Intelligible speech encoded in the human brain stem frequency-following response, NeuroReport 6 (1995) 2363–2367.

[7] G.C. Galbraith, S.P. Jhaveri, J. Kuo, Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition, Electroencephalogr. Clin. Neurophysiol. 102 (1997) 46–53.

[8] J. Gandour, Phonetics of tone, in: R. Asher, J. Simpson (Eds.), The Encyclopedia of Language and Linguistics, vol. 6, Pergamon Press, New York, 1994, pp. 3116–3123.

[9] J. Gandour, Aphasia in tone languages, in: P. Coppens, A. Basso, Y. Lebrun (Eds.), Aphasia in Atypical Populations, Lawrence Erlbaum, Hillsdale, NJ, 1998, pp. 117–141.

[10] J. Gandour, D. Wong, L. Hsieh, B. Weinzapfel, D. Van Lancker, G.D. Hutchins, A crosslinguistic PET study of tone perception, J. Cogn. Neurosci. 12 (2000) 207–222.

[11] E.M. Glaser, C.M. Suter, R. Dasheiff, A. Goldberg, The human frequency-following response: its behavior during continuous tone and tone burst stimulation, Electroencephalogr. Clin. Neurophysiol. 40 (1976) 25–32.

[12] J.I. Gold, E.I. Knudsen, A site of auditory experience-dependent plasticity in the neural representation of auditory space in the barn owl's inferior colliculus, J. Neurosci. 20 (2000) 3469–3486.

[13] S. Greenberg, J.T. Marsh, W.S. Brown, J.C. Smith, Neural temporal coding of low pitch: I. Human frequency-following responses to complex tones, Hear. Res. 25 (1987) 91–114.

[14] J.W. Hall III, Auditory brainstem frequency following responses to waveform envelope periodicity, Science 205 (1979) 1297–1299.

[15] G. Hickok, D. Poeppel, Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language, Cognition 92 (2004) 67–99.

[16] J.M. Howie, Acoustical Studies of Mandarin Vowels and Tones, Cambridge University Press, New York, 1976.

[17] L. Hsieh, J. Gandour, D. Wong, G.D. Hutchins, Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience, Brain Lang. 76 (2001) 227–252.

[18] D.H. Klatt, Software for a cascade/parallel formant synthesizer, J. Acoust. Soc. Am. 67 (1980) 971–995.

[19] E.I. Knudsen, W. Zheng, W.M. DeBello, Traces of learning in the auditory localization pathway, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 11815–11820.

[20] A. Krishnan, Human frequency-following responses to two-tone approximations of steady-state vowels, Audiol. Neuro-otol. 4 (1999) 95–103.

[21] A. Krishnan, Human frequency-following responses: representation of steady-state synthetic vowels, Hear. Res. 166 (2002) 192–201.

[22] A. Krishnan, J. Parkinson, Human frequency-following response: representation of tonal sweeps, Audiol. Neuro-otol. 5 (2000) 312–321.

[23] A. Krishnan, Y. Xu, J.T. Gandour, P.A. Cariani, Human frequency-following response: representation of pitch contours in Chinese tones, Hear. Res. 189 (2004) 1–12.

[24] J.T. Marsh, W.S. Brown, J.C. Smith, Differential brainstem pathways for the conduction of auditory frequency-following responses, Electroencephalogr. Clin. Neurophysiol. 36 (1974) 415–424.

[25] M.I. Miller, M.B. Sachs, Representation of voice pitch in discharge patterns of auditory-nerve fibers, Hear. Res. 14 (1984) 257–279.

[26] A.R. Palmer, I.M. Winter, C.J. Darwin, The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons, J. Acoust. Soc. Am. 79 (1986) 100–113.

[27] B. Philibert, L. Collet, E. Veuillet, Electrophysiological changes following hearing aid fitting in hearing-impaired listeners: evidence of a functional plasticity. Hear Res (in press).

[28] P.N. Plyler, A.K. Ananthanarayan, Human frequency-following responses: representation of second formant transitions in normal-hearing and hearing-impaired listeners, J. Am. Acad. Audiol. 12 (2001) 523–533.

[29] N.M. Russo, T.G. Nicol, S.G. Zecker, E.A. Hayes, N. Kraus, Auditory training improves neural timing in the human brainstem, Behav. Brain Res. 156 (2005) 95–103.

[30] R. Salmelin, A. Schnitzler, L. Parkkonen, K. Biermann, P. Helenius, K. Kiviniemi, K. Kuukka, F. Schmitz, H. Freund, Native language, gender, and functional organization of the auditory cortex, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 10460–10465.

[31] D.A. Schwartz, D. Purves, Pitch is determined by naturally occurring periodic sounds, Hear. Res. 194 (2004) 31–46.

[32] J.C. Smith, J.T. Marsh, W.S. Brown, Far-field recorded frequency-following responses: evidence for the locus of brainstem sources, Electroencephalogr. Clin. Neurophysiol. 39 (1975) 465–472.

[33] N. Suga, Biosonar and neural computation in bats, Sci. Am. 262 (1990) 60–68.

[34] N. Suga, Processing of auditory information carried by complex species-specific sounds, in: M.S. Gazzaniga, E. Bizzi (Eds.), The Cognitive Neurosciences, MIT Press, Cambridge, MA, 1994, pp. 295–318.

[35] N. Suga, E. Gao, Y. Zhang, X. Ma, J.F. Olsen, The corticofugal system for hearing: recent progress, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 11807–11814.

[36] N. Suga, X. Ma, E. Gao, M. Sakai, S.A. Chowdhury, Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists, Speech Commun. 41 (2003) 189–200.

[37] M. Vihla, K. Kiviniemi, R. Salmelin, Auditory cortical activation in Finnish and Swedish speaking Finns: a magnetoencephalographic study, Neurosci. Lett. 322 (2002) 141–144.

[38] Y. Wang, D.M. Behne, A. Jongman, J.A. Sereno, The role of linguistic experience in the hemispheric processing of lexical tone, Appl. Psycholinguist. 25 (2004) 449–466.

[39] Y. Xu, Contextual tonal variations in Mandarin, J. Phon. 25 (1997) 61–83.

[40] J. Yan, Y. Zhang, G. Ehret, Corticofugal shaping of frequency tuning curves in the central nucleus of the inferior colliculus of mice, J. Neurophysiol. 93 (2005) 71–83.

[41] R.J. Zatorre, P. Belin, V.B. Penhune, Structure and function of auditory cortex: music and speech, Trends Cogn. Sci. 6 (2002) 37–46.