Linguistic status of timbre influences pitch encoding in the brainstem

Ananthanarayan Krishnan[a], Jackson T. Gandour[a], Saradha Ananthakrishnan[a], Gavin M. Bidelman[a] and Christopher J. Smalt[b]

The aim of this experiment is to assess the effects of the linguistic status of timbre on pitch processing in the brainstem. Brainstem frequency following responses were evoked by the Mandarin high-rising lexical tone superimposed on a native vowel quality ([i]), nonnative vowel quality ([œ]), and iterated rippled noise (nonspeech). Results revealed that voice fundamental frequency magnitudes were larger when concomitant with a native vowel quality compared with either nonnative vowel quality or nonspeech timbre. Such experience-dependent effects suggest that subcortical sensory encoding of pitch interacts with timbre in the human brainstem. As a consequence, responses of the perceptual system can be differentially shaped to pitch patterns in relation to the linguistic status of their concomitant timbre. ***NeuroReport*** **22:801–803 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

NeuroReport 2011, **22**:801–803

Keywords: auditory, experience-dependent plasticity, frequency following response, human, iterated rippled noise, Mandarin Chinese, pitch, speech, timbre, vowel quality

[a]Department of Speech Language Hearing Sciences and [b]School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, USA

Correspondence to Professor Jackson T. Gandour, PhD, Department of Speech Language Hearing Sciences, 1353 Heavilon Hall, 500 Oval Drive, Purdue University, West Lafayette, IN 47907–2038, USA
Tel: +1 765 494 3821; fax: +1 765 494 0771;
e-mail: gandour@purdue.edu

Present address: Gavin M. Bidelman, Rotman Research Institute, Toronto, Ontario, Canada.

Received 12 July 2011 **accepted** 20 July 2011

It is now well documented that music and language experience enhances neural representation of information relevant to pitch and timbre at the level of the brainstem well before the auditory signal reaches the cerebral cortex [1,2]. On the basis of studies reporting an enhancement of neural representation of specific elements of pitch in the auditory brainstem responses of musicians [3] and tone language speakers [4], we infer that early auditory processing is subject to neural plasticity that manifests itself in stimuli that contains perceptually salient acoustic features, which occur within the listener's domain of expertise. The question of the degree of specificity in experience-dependent brainstem representation, however, is one that warrants further empirical investigation.

Of particular interest here is whether manipulation of the spectral components (i.e., timbre) of a complex sound will influence the representation of pitch in the auditory brainstem of individuals who are native speakers of Mandarin. It is clear from psychoacoustic studies that manipulation of spectral components (resolved/unresolved harmonics) influences the discriminability and salience of pitch [5]. In addition, empirical studies indicate that the heard pitch of speech [6] or music [7] is dependent on timbre. Indeed, pianists' responses, for example, represent the timbral characteristics of piano sounds with greater fidelity than those of nonpianists [8]. Thus, in the music domain, the brainstem is sensitive to individuals' long-term exposure to specific timbres. Given these findings, we reasoned that experience-dependent enhancement of pitch-relevant information may be modulated by timbral characteristics of speech in the language domain.

Analogs to a pianist listening to music fixed in pitch, but differing in instrumental quality, we asked native Chinese to listen to a lexical tone (Mandarin tone 2, T2), fixed in linguistic pitch, but differing in vowel quality (cf. timbre). The native tone was superimposed on a native vowel quality (high-front unrounded, [i]), nonnative vowel quality (low front unrounded, [œ]), and nonspeech timbre [iterated rippled noise (IRN)]. Accordingly, we expected their brainstem responses to the pitch of a Mandarin tone to vary depending on the linguistic status of its concomitant timbre.

Methods

Participants

Eleven right-handed native speakers (six male participants; mean age, 24.5 years) of Mandarin Chinese participated in the experiment. They demonstrated normal hearing audiometric thresholds (0.5–4 kHz), and gave informed consent in compliance with the Purdue University Institutional Review Board. All had less than 3 years of musical training.

Stimuli

Two pairs of stimuli ($T2_i$/$T2_{oe}$, native/nonnative; $T2_i$/$T2_{irn}$, speech/nonspeech) were constructed that varied only in their timbre. Vowel qualities were generated using a formant synthesizer [9]. Formant values were (in Hz): [i] F1(300), F2(2500), F3(3500), F4(4530); [œ] F1(465), F2(1186), F3(2281), F4(3153). For $T2_{irn}$,

0959-4965 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins

DOI: 10.1097/WNR.0b013e32834b2996

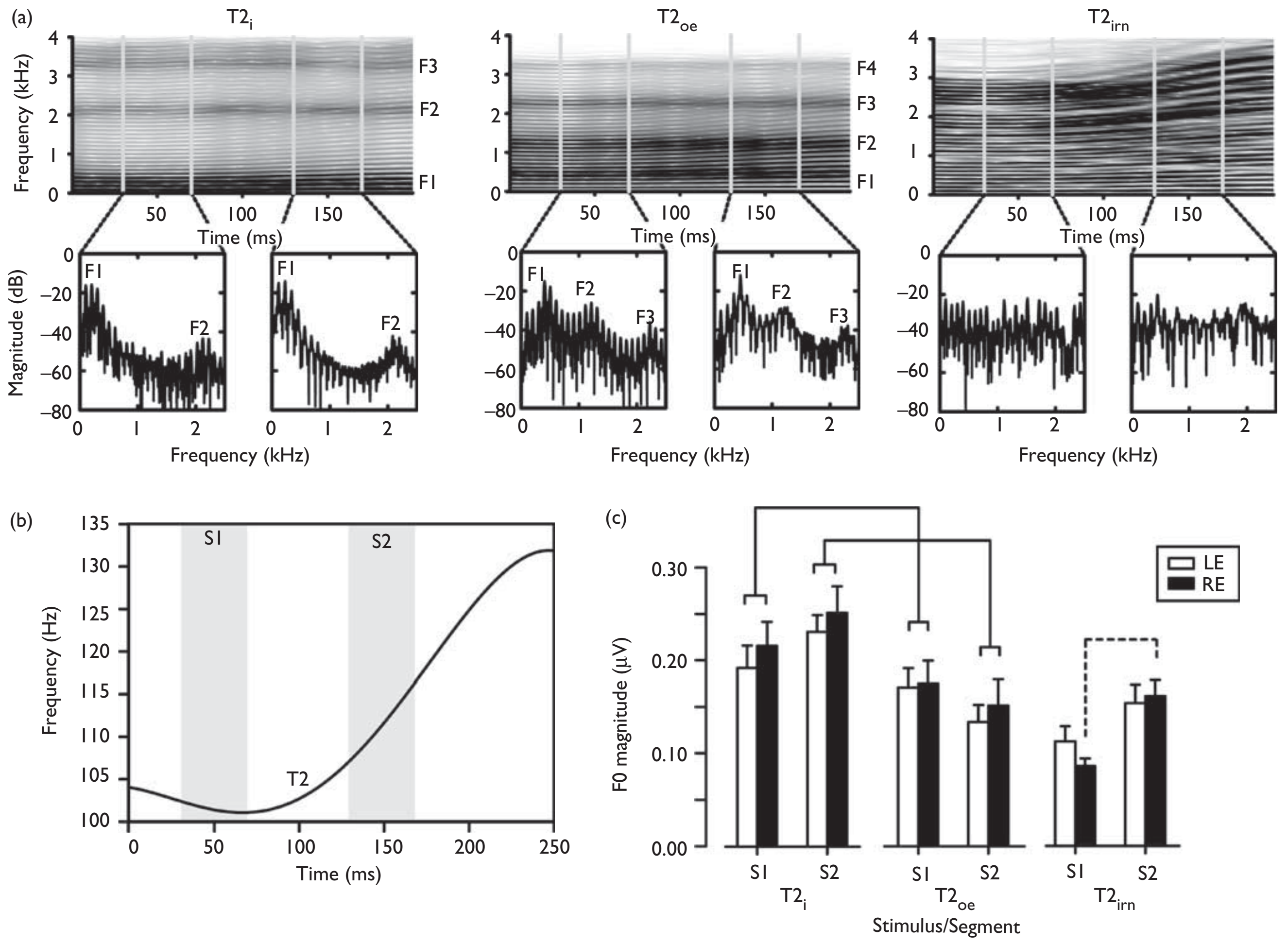
formant structure was removed, thus producing a timbre uncharacteristic of natural speech [10] (Fig. 1a). An F0 pitch sweep, modeled after a natural production of T2 was superimposed on all three stimuli (Fig. 1b). Stimuli were normalized in both duration (250 ms) and overall RMS amplitude. The f_0 contour of T2 was modeled after its natural citation form, as produced by a male speaker [11], using a fourth-order polynomial [12].

Data acquisition

The frequency following response (FFR) recording protocol was identical to that used in Krishnan *et al.* [13]. FFRs were recorded from each participant in response to monaural stimulation of the left ear (LE) and the right ear (RE) through magnetically shielded insert earphones (Etymotic Research-3A) at 80 dB of sound pressure level (rarefaction polarity; 2.43/s repetition rate). Stimulus presentation and data acquisition were accomplished using the SmartEP software within the evoked potential system (Intelligent Hearing System, Miami, USA).

FFRs were recorded differentially between Ag-AgCl scalp electrodes placed on the midline of the forehead at the hairline (approximately Fpz) and the right mastoid (A2) or left mastoid (A1). Another electrode placed on the mid forehead served as a common ground. The raw electroencephalogram was amplified by 200 000 and filtered online (30–5000 Hz). Inter-electrode impedances were maintained at less than or equal to 1 kΩ. Individual sweeps were recorded using an analysis window of 280 ms at a sampling rate of 10 kHz. Neural responses were further band pass filtered offline (80–2500 Hz). Sweeps containing activity exceeding ± 40 μV were rejected as artifacts. In total, each FFR waveform represents the average of 3000 artifact-free stimulus presentations.

Fig. 1



(a) Narrowband spectrograms of stimuli ($T2_i$/$T2_{oe}$/$T2_{irn}$) and spectra of time segments of analysis centered at 50 and 150 ms. (b) Fundamental frequency contour of T2. (c) Comparisons of pitch encoding show that a native vowel ($T2_i$) evokes a larger brainstem response than nonnative ($T2_{oe}$) in both ears across time segments (solid lines). Similarly, speech ($T2_i$) evokes a larger frequency following response (FFR) than nonspeech timbre ($T2_{irn}$). Using iterated rippled noise (IRN), a native pitch contour ($T2_{irn}$) evokes a larger right ear FFR in a perceptually salient portion (S2) of T2. F1/F2/F3 mark locations of formant peaks. S1/S2 represent time segments of analysis; T2, Mandarin tone 2.

Data analysis

FFR pitch encoding was quantified by measuring the spectral magnitude of the FFR component at F0 from each response waveform for each stimulus per ear. Two 40-ms segments (S1: 30–70; S2: 130–170) were extracted from each FFR. S2 was chosen, as it coincides with those portions of T2 that contribute importantly to tonal recognition [14] and brainstem representation [3,15]. For each condition per segment, the magnitude of F0 was measured as the peak in the FFT, relative to the noise floor, between 100 and 135 Hz. All FFR analyses were performed. All data analyses were performed using custom routines coded in Matlab 7.10 (The MathWorks, Inc., Natick, Massachusetts, USA).

Statistical analysis

Per stimulus pair ($T2_i/T2_{oe}$; $T2_i/T2_{irn}$), two-way, mixed-model analyses of variance were conducted on ears separately to assess the effects of timbre experience and time segments on pitch encoding. Bonferroni corrections were applied to multiple pairwise comparisons.

Results

Analyses of variance revealed that F0 magnitudes were larger in both ears across time segments when concomitant with a native vowel quality, $T2_i$, compared with a non-native vowel quality, $T2_{oe}$, [LE: $F(1,10)=8.39$, $P=0.0159$; RE: $F=6.89$, $P=0.0254$], or nonspeech timbre, $T2_{irn}$, [LE: $F(1,10)=15.91$, $P=0.0026$; RE: $F=26.17$, $P=0.0005$] (Fig. 1c). Within $T2_{irn}$, a time segment effect in the RE showed that the F0 magnitude of S2 was larger compared with S1 ($F=6.59$, $P=0.0280$).

Discussion

Psychoacoustically, manipulation of spectral components (resolved/unresolved harmonics) influences the discriminability and salience of pitch [5]. Indeed, empirical studies indicate that the heard pitch of speech [6] or music [7] is dependent on timbre. Here, we have shown that the linguistic status of timbre may influence brainstem encoding of linguistic pitch. FFRs evoked by the fully native stimulus ($T2_i$) were larger than stimuli deviant in timbre ($T2_{oe}$, $T2_{irn}$). This complementary interaction between pitch and timbre, when both the source (pitch) and filter (spectral) characteristics are native, suggests that nascent representations of acoustic–phonetic features emerge early along the auditory pathway. Just as pianists are tuned more closely to musical pitch played on a piano [8], so too are Chinese tuned more closely to linguistic pitch played on a native vowel. Such effects, however, are not likely to be restricted to Mandarin. Indeed, enhancement of brainstem pitch encoding may transfer to other languages or domains (e.g., music→language) as long as they share similar features, which are of perceptual significance to the listener [3,4]. Within $T2_{irn}$, FFRs elicited from the RE are larger in a perceptually salient portion (S2) relative to S1 [14]. It is unlikely that differences in F0 magnitude between the two segments (S1, S2) can explain our results because $T2_{irn}$ actually showed a 2.1 dB greater F0 spectral magnitude in S1 when compared with S2. In Mandarin, vowels exert greater interference on tones than *vice versa* [16]. The perceptual saliency of S2 in the RE may mirror engagement of the left hemisphere in processing lexical tones. Although our data cannot distinguish neural mechanisms involved in top–down modulation from those local to the brainstem, the corticofugal system is known to mediate the learning of behaviorally relevant auditory features [17].

Acknowledgements

Sources of support: National Institutes of Health R01 DC008549 (A.K.); Bilsland dissertation fellowship (G.M.B.).

Conflicts of interest

There are no conflicts of interest.

References

1 Krishnan A, Bidelman GM, Gandour JT. Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hear Res* 2010; **268**:60–66.

2 Kraus N, Banai K. Auditory-processing malleability: focus on language and music. *Curr Direct Psychol Sci* 2007; **16**:105–110.

3 Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 2011; **23**:425–434.

4 Krishnan A, Gandour JT, Bidelman GM. The effects of tone language experience on pitch processing in the brainstem. *J Neurolinguistics* 2010; **23**:81–95.

5 Ritsma RJ. Frequencies dominant in the perception of the pitch of complex sounds. *J Acoust Soc Am* 1967; **42**:191–198.

6 Stoll G. Pitch of vowels: experimental and theoretical investigation of its dependence on vowel quality. *Sp Comm* 1984; **3**:137–150.

7 Krumhansl CL, Iverson P. Perceptual interactions between musical pitch and timbre. *J Exp Psychol Hum Percep Perform* 1992; **18**:739–751.

8 Strait DL, Chan K, Ashley R, Kraus N. Specialization among the specialized: auditory brainstem function is tuned in to timbre. *Cortex* 2011 doi: 10.1016/j.cortex.2011.03.015.

9 Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am* 1990; **87**:820–857.

10 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 Trans Biomed Eng* 2008; **55**:281–287.

11 Xu Y. Contextual tonal variations in Mandarin. *J Phon* 1997; **25**:61–83.

12 Xu Y, Gandour J, Talavage T, Wong D, Dzemidzic M, Tong Y, *et al*. Activation of the left planum temporale in pitch processing is shaped by language experience. *Hum Brain Mapp* 2006; **27**:173–183.

13 Krishnan A, Gandour JT, Ananthanarayan AK, Bidelman GM, Smalt CJ. Functional ear (a)symmetry in brainstem neural activity relevant to encoding of voice pitch: a precursor for hemispheric specialization? *Brain Lang* (in press). doi: 10.1016/j.bandl.2011.05.001.

14 Whalen DH, Xu Y. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 1992; **49**:25–47.

15 Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J Cogn Neurosci* 2009; **21**:1092–1105.

16 Tong Y, Francis A, Gandour JT. Processing dependencies between segmental and suprasegmental features of Mandarin Chinese. *Lang Cognit Process* 2008; **23**:689–708.

17 Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat Neurosci* 2010; **13**:253–260.