Published in final edited form as:

Neuroreport. 2011 November 16; 22(16): 801–803. doi:10.1097/WNR.0b013e32834b2996.

Linguistic status of timbre influences pitch encoding in the brainstem

Ananthanarayan Krishnan^a, Jackson T. Gandour^a, Saradha Ananthakrishnan^a, Gavin M. Bidelman^{a,c}, and Christopher J. Smalt^b

Ananthanarayan Krishnan: rkrish@purdue.edu; Jackson T. Gandour: gandour@purdue.edu; Saradha Ananthakrishnan: sanantha@purdue.edu; Gavin M. Bidelman: gbidelman@rotman-baycrest.on.ca; Christopher J. Smalt: csmalt@purdue.edu

^aDepartment of Speech Language Hearing Sciences, Purdue University

^bSchool of Electrical and Computer Engineering, Purdue University

Abstract

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 (non-speech). Results revealed that voice fundamental frequency magnitudes were larger when concomitant with a native vowel quality as compared to either nonnative vowel quality or non-speech 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.

Keywords

auditory; human; speech; pitch; timbre; vowel quality; iterated rippled noise (IRN); frequency following response (FFR); Mandarin Chinese; experience-dependent plasticity

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]. Based on studies reporting 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 contain 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) influence the

Corresponding author: Jackson T. Gandour, PhD, Purdue University, Department of Speech Language Hearing Sciences, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038 USA, TEL: 765.494.3821, FAX: 765.494.0771, gandour@purdue.edu.

^cPresent address: Rotman Research Institute, Toronto, ON, Canada

Conflicts of Interest: none declared

discriminability and salience of pitch [5]. Also, 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 non-pianists [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.

Analogous 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, [æ]), and nonspeech timbre (iterated rippled noise, IRN). Accordingly, we expected their brainstem responses to the pitch of a Mandarin tone to vary depending upon the linguistic status of its concomitant timbre.

Methods

Subjects

Eleven right-handed native speakers (6 male; mean age 24.5) of Mandarin Chinese participated in the experiment. Subjects 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 (T2i/T2oe, native/nonnative; T2i/T2irn, 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); [æ] F1(465), F2(1186), F3(2281), F4(3153). For T2irn, 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 msec) 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 FFR recording protocol was identical to that used in Krishnan et al. [13]. FFRs were recorded from each subject in response to monaural stimulation of the left (LE) and right ear (RE) through magnetically shielded insert earphones (ER-3A) at 80 dB SPL (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 (~Fpz) and the right mastoid (A2) or left mastoid (A1). Another electrode placed on the mid-forehead served as common ground. The raw EEG was amplified by 200,000 and filtered online (30–5000 Hz). Inter-electrode impedances were maintained ≤ 1 k Ω . 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 μ V were rejected as artifacts. In total, each FFR waveform represents the average of 3000 artifact-free stimulus presentations.

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-msec segments (S1: 30–70; S2: 130–170) were extracted from each FFR. S2 was chosen because 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 100–135 Hz. All FFR analyses were performed. All data analyses were performed using custom routines coded in MATLAB ®7.10 (The MathWorks, Inc., Natick, MA).

Statistical analysis

Per stimulus pair (T2i/T2oe; T2i/T2irn), two-way, mixed-model ANOVAs 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

ANOVAs revealed that F0 magnitudes were larger in both ears across time segments when concomitant with a native vowel quality, T2i, as compared to a nonnative vowel quality, T2oe, [LE: $F_{1,10} = 8.39$, $p = 0.0159$; RE: $F = 6.89$, $p = 0.0254$], or nonspeech timbre, T2irn, [LE: $F_{1,10} = 15.91$, $p = 0.0026$; RE: $F = 26.17$, $p = 0.0005$] (Fig. 1C). Within T2irn, a time segment effect in the RE showed that the F0 magnitude of S2 was larger than that of S1 [$F = 6.59$, $p = 0.0280$].

Discussion

Psychoacoustically, manipulation of spectral components (resolved/unresolved harmonics) influence 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 (T2i) were larger than stimuli deviant in timbre (T2oe, T2irn). 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 T2irn, 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 T2irn actually showed a 2.1 dB greater F0 spectral magnitude in S1 when compared to 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. Though 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].

Acknowledgments

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

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. [PubMed: 20457239]
2. Kraus N, Banai K. Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science.* 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. [PubMed: 19925180]
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. [PubMed: 20161561]
5. Ritsma RJ. Frequencies dominant in the perception of the pitch of complex sounds. *J Acoust Soc Am.* 1967; 42:191–198. [PubMed: 6052077]
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. *Journal of experimental psychology Human perception and performance.* 1992; 18:739–751. [PubMed: 1500873]
8. Strait DL, Chan K, Ashley R, Kraus N. Specialization among the specialized: Auditory brainstem function is tuned in to timbre. *Cortex.* 2011
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. [PubMed: 2137837]
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. [PubMed: 18232372]
11. Xu Y. Contextual tonal variations in Mandarin. *J Phon.* 1997; 25:61–83.
12. Xu Y, Gandour J, Talavage T, Wong D, Dziedzic 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. [PubMed: 16035045]
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. 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. [PubMed: 1603839]
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. [PubMed: 18702588]
16. Tong Y, Francis A, Gandour JT. Processing dependencies between segmental and suprasegmental features of Mandarin Chinese. *Language and Cognitive Processes.* 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. [PubMed: 20037578]

Figure 1.

(A) Narrowband spectrograms of stimuli (T2_i/T2_{oe}/T2_{irn}) and spectra of time segments of analysis centered at 50 and 150 msec. (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 FFR than nonspeech timbre (T2_{irn}). Using 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; FFR, frequency following response; IRN, iterated rippled noise.