Brain & Language 114 (2010) 193–198

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Short Communication

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

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

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

ARTICLE INFO

Article history:
Accepted 12 May 2010
Available online 8 June 2010

Keywords:
Auditory
Human
Brainstem
Pitch
Language
Frequency-following response (FFR)
Click trains
Mandarin Chinese
Experience-dependent plasticity
Speech perception

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.

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.

* Corresponding author. Address: Department of Speech Language Hearing Sciences, Purdue University, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038, USA. Fax: +1 765 494 0771.
E-mail addresses: rkrish@purdue.edu (A. Krishnan), gandour@purdue.edu (J.T. Gandour).

0093-934X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2010.05.004

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 shows more robust representation of the first five 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 × 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