

Experience-dependent neural plasticity is sensitive to shape of pitch contours

Bharath Chandrasekaran, Ananthanarayan Krishnan and Jackson T. Gandour

Department of Speech Language Hearing Sciences, Purdue University, West Lafayette, Indiana, USA

Correspondence to Jackson T. Gandour, PhD, Department of Speech Language Hearing Sciences, Purdue University, 1353 Heavilon Hall, 500 Oval Drive, West Lafayette, IN 47907-2038, USA

Tel: +1 765 494 3821; fax: +1 765 494 0771; e-mail: gandour@purdue.edu

Author to whom requests for reprints should be made: Ananthanarayan Krishnan (e-mail: rkrish@purdue.edu)

Received 7 August 2007; accepted 10 September 2007

Language experience is known to modulate the preattentive processing of linguistically relevant pitch contours when presented in the speech domain. To assess if experience-dependent effects are specific to speech, we evaluated the mismatch negativity response to nonspeech homologs (iterated rippled noise) of such curvilinear pitch contours (Mandarin: Tone 1, 'high level'; Tone 2, 'high rising') by Chinese and English listeners as well as to a pitch contour that was a linear approximation of Tone 2 ('linear ascending

ramp'). Mandarin speakers showed larger mismatch negativity responses than English to the curvilinear pitch contours only. These results suggest that experience-dependent neural plasticity in early cortical processing of linguistically relevant pitch contours is sensitive to naturally occurring pitch dimensions but not specific to speech *per se*. *NeuroReport* 18:1963–1967 © 2007 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: auditory, experience-dependent plasticity, iterated rippled noise, lexical tone, Mandarin Chinese, mismatch negativity, pitch

Introduction

Languages that exploit phonologically contrasting variations in pitch at the word or syllable level are called tone languages [1]. Mandarin Chinese is a tone language. In addition to consonants and vowels, Chinese has four tones: ma¹ 'mother' [T1], ma² 'hemp' [T2], ma³ 'horse' [T3], ma⁴ 'scold' [T4]. Tones 1–4 can be described phonetically as high level, high rising, low falling rising, and high falling, respectively [2].

Using speech stimuli representative of Mandarin lexical tones in a passive oddball paradigm [3,4], it has been demonstrated that the T1/T2 and T1/T3 conditions, which involve a contrast between level (T1) and contour (T2, T3) tones, elicit larger mismatch negativity (MMN) in the Chinese group relative to the English. Thus, we infer that automatic, involuntary, preattentive processing of lexical tones at early stages of pitch processing in the cortex may be shaped by a listeners' long-term familiarity with the pitch contours of a particular language. A multidimensional scaling analysis of MMN responses further reveals that the Chinese group is more sensitive to pitch contour than the English group [3]. Thus, MMN indexes the relevance of pitch contour in early cortical stages of tonal processing. The question arises whether this experience-dependent neural plasticity is speech-specific. Whereas differences in MMN responses may emerge from language experience, the effects of such experience are not necessarily specific to speech perception.

By using iterated rippled noise (IRN) stimuli to elicit the MMN, we can now present linguistically relevant pitch contours in a nonspeech context, thereby allowing us to

evaluate whether experience-dependent plasticity to pitch contours is domain-general. An IRN stimulus is created by repeatedly delaying and adding broadband noise to itself. The perceived pitch has been shown to correspond with a reciprocal of delay; pitch salience increases as a function of the number of iterations [5,6]. The IRN algorithm has recently been modified to incorporate multiple time-dependent delays over a range of iteration steps thereby allowing for dynamic changes in pitch [7] and especially curvilinear pitch contours that occur in natural speech [8]. These modifications enable us to investigate the neural processing of linguistically relevant pitch contours in the nonspeech domain. IRN stimuli eliminate any lexical-semantic confound inherent to speech. Thus, any advantage in the processing of IRN homologs of lexical tones for native relative to non-native speakers cannot be due to psycholinguistic factors.

The aim was to demonstrate that language-dependent neural plasticity in the processing of pitch in the cortex, as reflected by the MMN, is not specific to speech. Two IRN stimuli were selected to represent prototypical, curvilinear Mandarin tones (T1, T2). A third IRN stimulus (T2L) was selected to represent a linear rising ramp that does not occur in Mandarin citation form or running speech. By adding T2L, we were able to test whether experience-dependent plasticity at early cortical stages of pitch processing is sensitive to contours that naturally occur in Mandarin speech. By including two language groups, one native (Chinese), the other non-native (English), we were able to determine whether modulation of the MMN in response to nonspeech

stimuli is due to a listener's long-term familiarity with specific pitch dimensions. If not, then we would expect uniform MMN responses for both oddball conditions regardless of a listener's language experience. One condition was composed of T1 and T2, both of which occur in natural speech. The other condition involved T1 and T2L, the latter of which exhibited a dimension not observed in natural speech. By hypothesis, we expected the Chinese group to show larger MMN amplitude relative to the English group in response to the T1/T2 condition, whereas no group differences were expected in response to the T1/T2L condition.

Methods

Participants

Ten adult native speakers of Mandarin (four men, six women) and 10 adult native speakers of American English (four men, six women) took part in the experiment. The two groups were closely matched in age ($M \pm SD$: Chinese = 28.2 ± 2.3 ; English = 27.5 ± 3.2) and education ($M \pm SD$: Chinese = 18.2 ± 2.2 ; English = 17.1 ± 2.8). They were strongly right handed ($\geq 95\%$) as measured by the Edinburgh Handedness Inventory [9]. All participants exhibited normal hearing sensitivity (20 dB HL) at frequencies of 0.5, 1, 2, and 4 kHz. All participants completed a language history questionnaire [10]. Native speakers of Mandarin Chinese did not have any English instruction before the age of 11 years. The English group had no prior experience learning any tonal language. All participants were non-musicians as determined by a music history questionnaire [11]. None of the native Mandarin or American English speakers had more than 3 years of formal training in music or any combination of musical instruments, and none had any type of musical training within the past 5 years. All participants gave their informed consent in compliance with the protocol approved by the institutional review board of Purdue University.

Stimuli

IRN was used to create three time-varying nonspeech f_0 contours using procedures similar to those described in [8]. A high iteration step (32) was used for all contrasts with gain set to 1. Of the three time-varying f_0 contours (Fig. 1), two (T1, T2) were modeled after natural citation-form Mandarin f_0 contours using a fourth-order polynomial

equation [12]. T1 and T2 reflect native Mandarin high level and rising tones, respectively, differing from each other on the basis of onset, offset, height, direction, and shape of f_0 contour. The third stimulus (T2L) was a linear approximation of T2, having the same onset, offset, and direction as T2. The duration of all three stimuli was fixed at 250 ms with 10 ms rise/fall time. Amplitude was fixed at 70 dB.

Data acquisition

Participants sat in an acoustically and electrically shielded booth facing an LCD monitor. They were instructed to ignore the sounds presented binaurally via insert earphones and watch a self-selected movie with subtitles. Participants were also informed that they would have to provide a synopsis of the movie at the end of the experimental session. The interstimulus onset-to-onset interval was fixed at 667 ms. For all oddball sequences, the frequent stimulus (standard) was presented at a probability of 0.85 and the infrequent stimulus (deviant) occurred at a probability of 0.15. Within the oddball sequences, the order of presentation of stimuli was pseudorandom, that is at least one standard stimulus preceded the deviant.

The experiment consisted of four oddball sequences. In one condition (T1/T2), T1 (curvilinear: level) was presented as the standard ($P=0.85$) and T2 (curvilinear: rising) as the deviant ($P=0.15$). In a second condition, T1 was presented as the standard, with T2L (linear: rising) as the deviant. In the two remaining sequences, the oddball sequences were reversed with T1 as the deviant ($P=0.15$) with T2 and T2L separately as standards ($P=0.85$). A hundred artifact-free deviants were collected for each sequence. The experiment ran for approximately 2 h including subject preparation. All stimuli were controlled by a signal generation and data acquisition system (Smart EP, Intelligent Hearing Systems Inc. Miami, Florida, USA). Stimuli were presented binaurally at 75 dB SPL to each ear through magnetically shielded insert earphones (Biologic TIP-300, Biologic Corporation, Madison, Wisconsin, USA).

For each participant, AgCl electrodes were mounted on the frontal midline (Fz), central midline (Cz) locations according to the 10–20 location system. These two electrode locations were chosen because the typical MMN response is known to be the most robust at the frontal electrode sites [13] and shows a distinct reduction in amplitude at more

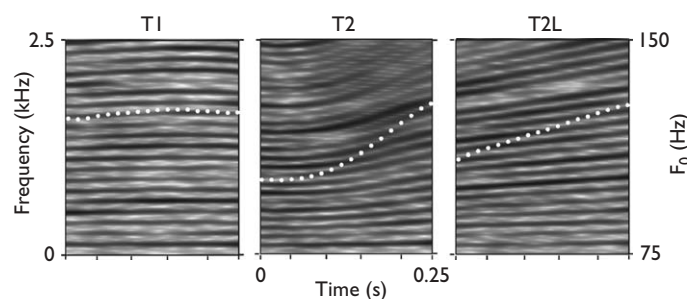


Fig. 1 Narrowband spectrograms of the three IRN stimuli used in this study. F_0 contours are plotted as white dots. T1 and T2 are modeled after average f_0 contours of time-normalized Mandarin lexical tones [25]. T2L, a linear rising ramp, represents a pitch contour that, albeit a crude approximation of T2, does not actually occur in natural speech. T1 (left panel) was standard in both conditions; T2 (middle panel), the curvilinear rising contour, served as the deviant in one condition (T1/T2); T2L (right panel), the linear rising ramp, served as the deviant in the other condition (T1/T2L). Onsets and offsets of T2 and T2L were identical. T1, Mandarin high level tone; T2, Mandarin high rising tone; T2L, linear rising ramp that does not occur in Mandarin tonal inventory. The three stimuli were created at a high iteration step from broadband noise [8]. Clear bands of energy are evident at the time-varying f_0 and its harmonics, but unlike speech, IRN stimuli show no formant structure. IRN, iterated rippled noise.

central sites. The tip of the nose served as the reference electrode, and the forehead served as the ground. The right and left mastoids were linked and used as a third reference site. The impedance across all electrodes was kept below 5 k Ω . Electrodes monitoring vertical eye movements were used to remove eyeblink-related artifacts, as defined by epochs with voltage changes exceeding 60 μ V. The signals were band-passed filtered at 1–30 Hz and recorded at a 1000 Hz sampling rate.

Data analysis

The baseline for the grand averaged waveforms was defined as the average of the amplitude values between –100 and 0 ms (onset of stimuli). In each experimental condition, the MMN was obtained by subtracting the response to the deviant in an oddball condition from the response to the same deviant presented as a standard in the reversed condition. This subtraction process, in which the deviant waveform is compared with the same stimuli presented as the standard, effectively controls for any acoustical differences between stimuli. The MMN peak latency was calculated as the most negative voltage in the MMN window between 125 and 300 ms. The MMN mean amplitude was calculated as the mean voltage from a 100 ms window centered on the MMN peak latency. MMN mean amplitudes and peak latencies were analyzed using a three-way mixed model analysis of variance (ANOVA) with subject as random effect, group (Chinese, English) as between subject factor, and condition (T1/T2, T1/T2L) and location (Fz, Cz) as within subject factors.

Results

Mismatch negativity mean amplitude

The grand average waveforms for the two groups (Chinese, English), two conditions (T1/T2, T1/T2L), across three locations (Fz, Cz, and linked mastoids) are shown in Fig. 2.

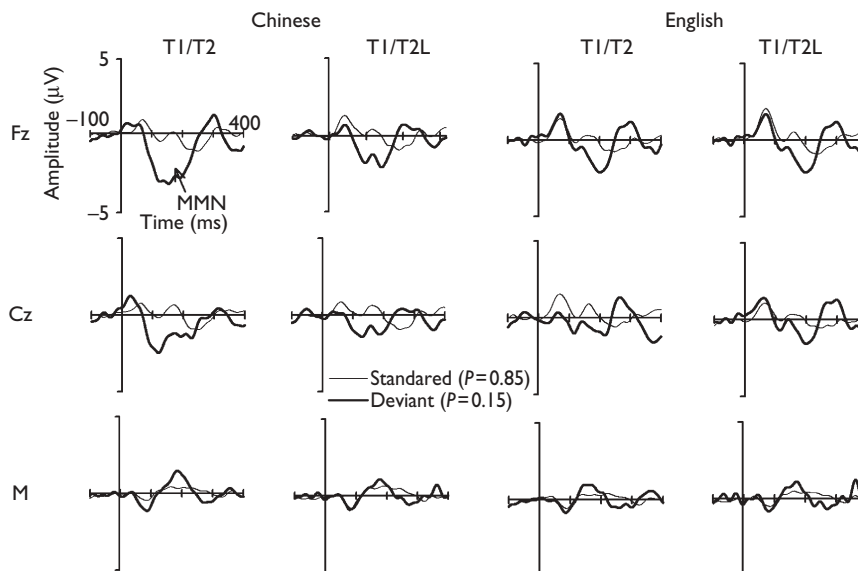


Fig. 2 Grand average standard ($P=0.85$) and deviant ($P=0.15$) waveforms displayed for the two language groups (Chinese, English) per experimental condition (T1/T2, T1/T2L) at the three electrode locations (Fz, Cz, mastoid). Irrespective of group or condition, the MMN peaked between 150 and 180 ms and was largest at the Fz electrode relative to the Cz electrode. In addition, the MMN showed the typical polarity reversal at the mastoid location. Grand average responses of the Chinese group showed a larger MMN-related negativity for the T1/T2 condition, relative to the English group. The two groups did not differ with respect to the T1/T2L condition. MMN, mismatch negativity.

Results from the omnibus three-way ANOVA (group-condition \times location) revealed significant main effects of group [$F(1,18)=24.86$; $P<0.0001$, $\eta^2_{\text{partial}}=0.58$], condition [$F(1,18)=7.12$; $P=0.016$, $\eta^2_{\text{partial}}=0.28$] and location [$F(1,36)=27.74$; $P<0.0001$, $\eta^2_{\text{partial}}=0.44$] and a significant interaction effect between group and condition [$F(1,18)=16.80$; $P=0.0007$, $\eta^2_{\text{partial}}=0.48$]. No significant interaction effects were observed between condition and location [$F(1,36)=3.15$; $P=0.08$], group and location [$F(1,36)=0.33$; $P=0.57$], or between group, condition, and location [$F(1,36)=0.16$; $P=0.69$]. MMN mean amplitude for each group (Chinese, English) and condition (T1/T2, T1/T2L) at the electrode location Fz are displayed in Fig. 3. Between-group comparisons revealed that Chinese participants had larger MMN mean amplitude relative to English for the T1/T2 condition [$F(1,18)=41.79$, $P<0.001$, $\eta^2_{\text{partial}}=0.70$], but not for the T1/T2L condition [$F(1,18)=1.23$; $P=0.28$]. Between-condition comparisons showed that MMN mean amplitude for the Chinese group was significantly less in the T1/T2L condition relative to the T1/T2 condition [$F(1,18)=22.81$; $P=0.0002$, $\eta^2_{\text{partial}}=0.56$]. In contrast, there was no significant difference between the two conditions for the English group [$F(1,18)=1.03$; $P=0.32$].

Mismatch negativity peak latency

For both groups, MMN peaked between 160 and 180 ms irrespective of condition, or location. ANOVA results did not yield significant main effects for group, condition, or location. None of the two-way or three-way interactions reached significance.

Discussion

The major finding of this crosslanguage study is that long-term experience with a tone language modulates pre-attentive cortical processing of nonspeech curvilinear pitch contours. Mandarin speakers exhibit larger MMN

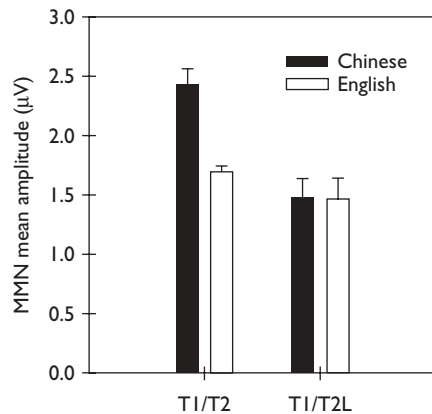


Fig. 3 Mean MMN amplitude for the two language groups (Chinese, English) per experimental condition (T1/T2, T1/T2L) as measured from the Fz electrode location. The mean MMN amplitude was larger in the Chinese group relative to the English group for the T1/T2 condition only. Only the Chinese group showed a larger response to the T1/T2 relative to the T1/T2L condition. MMN, mismatch negativity.

responses, as compared with English, in response to a deviant representing a curvilinear rising contour (T2) that occurs in natural speech, but not in response to a linear rising ramp (T2L) as a deviant. The fact that nonspeech homologs of native pitch contours (T1/T2) elicit enhanced MMN responses for native speakers of Mandarin relative to English suggest that experience-dependent neural plasticity is not speech-specific. Using IRN stimuli, this result is not confounded by lexical-semantic interference.

Rather, our data suggest that experience-dependent neural plasticity is feature or dimension-specific, that is language-dependent modulation of MMN occurs only for contrasts involving curvilinear pitch contours that occur in natural speech. Although T2L is similar in f_0 onset, offset, and direction as T2, it lacks the curvilinear shape of T2. As predicted, only the native Chinese group exhibits a larger MMN response to the T1/T2 in comparison to the T1/T2L condition.

These results are in agreement with a crosslanguage MMN study of another suprasegmental feature of speech-sound duration [14]. In Finnish, duration is phonemic (e.g. /tulli/ 'fire', /tuuli/ 'wind', /tulli/ 'customs'); in German, it is not. Native speakers of Finnish show enhanced MMN responses to nonspeech stimuli that differ in duration whereas German speakers do not. Whether frequency or duration, we infer that early preattentive cortical processing is selectively tuned to those features of the auditory signal that are relevant in a particular language even when presented in a nonspeech context.

A complete understanding of the neural organization of language can only be achieved within a framework involving a series of computations that apply to representations at different stages of processing [15]. A recent study suggests that the MMN may be more sensitive to subtle within-category acoustic differences than behavioral indices imply [16]. The linearity dimension (linear vs. curvilinear), of course, may be overridden at later stages of processing which recruit attention and memory components. Other behavioral data from multidimensional scaling [17] and categorical perception [18] of lexical tones reveal language-dependent effects even in response to linear f_0 ramps. At the level of the auditory brainstem, however, frequency follow-

ing responses elicited by curvilinear pitch contours (T1–T4) differ depending on listeners' language experience [19], whereas those elicited by linear rising (T2L) or falling (T4L) ramps are homogeneous regardless of language experience [20]. Our MMN data are compatible with the view that selective tuning of acoustic features relevant to speech begin at the earliest stages in the auditory pathway, but that speech-specific processing does not begin before the auditory signal extends beyond Heschl's gyrus [21]. A functional connectivity has been established between MMN and the timing of the brainstem onset response [22]. A similar connectivity is possible between MMN and frequency following responses. Our observed MMN responses may reflect a bottom-up influence from experience-driven, adaptive brainstem neural mechanisms [19,23].

Conclusion

Long-term experience with a tone language modulates early preattentive cortical processing of nonspeech curvilinear pitch contours. At this early stage, auditory processing is selectively tuned to linguistically relevant pitch dimensions that occur in natural speech. As indexed by the MMN, tuning is directed to pitch dimensions instead of tonal categories *per se*. These pitch dimensions are differentially weighted depending on a listener's experience with lexical tones. Minus the semantic confound of natural speech, the use of IRN stimuli enables us to investigate neural mechanisms underlying the processing of pitch contours that are 'linguistically relevant' comparable with those underlying the processing of 'behaviorally relevant' sounds in other nonprimate and nonhuman primate animals [24].

Acknowledgements

This study is based on part of a doctoral dissertation to be submitted by the first author at Purdue University in May 2008. B.C. is currently a predoctoral student in the Purdue University Life Sciences Integrative Neuroscience Program. The authors thank Jayaganesh Swaminathan for his assistance in generating the IRN stimuli; Bruce Craig and Eunjung Lim for their help with statistical analysis. Sources of support: NIH research Grant R01 DC008549-01 (A.K.); College of Liberal Arts (A.K. and J.G.).

References

1. Yip M. *Tone*. New York: Cambridge University Press; 2003.
2. Howie JM. *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press; 1976.
3. Chandrasekaran B, Gandour JT, Krishnan A. Neuroplasticity in the processing of pitch dimensions: a multidimensional scaling analysis of the mismatch negativity. *Restor Neurol Neurosci* 2007; **25**:195–210.
4. Chandrasekaran B, Krishnan A, Gandour JT. Mismatch negativity to pitch contours is influenced by language experience. *Brain Res* 2007; **1128**:148–156.
5. Patterson RD, Handel S, Yost WA, Datta AJ. The relative strength of the tone and noise components in iterated ripple noise. *J Acoust Soc Am* 1996; **100**:3286–3294.
6. Yost WA. Pitch strength of iterated ripple noise. *J Acoust Soc Am* 1996; **100**:3329–3335.
7. Denham S. Pitch detection of dynamic iterated rippled noise by humans and a modified auditory model. *Biosystems* 2005; **79**:199–206.
8. Swaminathan J, Krishnan A, Gandour JT. Applications of static and dynamic iterated rippled noise to evaluate pitch encoding in the human auditory brainstem. *IEEE Trans Biomed Eng* (in press).

9. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; **9**:97–113.
10. Li P, Sepanski S, Zhao X. Language history questionnaire: a web-based interface for bilingual research. *Behav Res Methods* 2006; **38**:202–210.
11. Wong PC, Perrachione TK. Learning pitch patterns in lexical identification by native English-speaking adults. *Appl Psycholinguist* 2007; **28**:565–585.
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. Naatanen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, et al. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 1997; **385**:432–434.
14. Tervaniemi M, Jacobsen T, Rottger S, Kujala T, Widmann A, Vainio M, et al. Selective tuning of cortical sound-feature processing by language experience. *Eur J Neurosci* 2006; **23**:2538–2541.
15. Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 2004; **92**:67–99.
16. Joanisse MF, Robertson EK, Newman RL. Mismatch negativity reflects sensory and phonetic speech processing. *NeuroReport* 2007; **18**:901–905.
17. Gandour J. Tone perception in Far Eastern languages. *J Phon* 1983; **11**:149–175.
18. Xu Y, Gandour JT, Francis AL. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *J Acoust Soc Am* 2006; **120**:1063–1074.
19. Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 2005; **25**:161–168.
20. Xu Y, Krishnan A, Gandour JT. Specificity of experience-dependent pitch representation in the brainstem. *NeuroReport* 2006; **17**:1601–1605.
21. Uppenkamp S, Johnsrude IS, Norris D, Marslen-Wilson W, Patterson RD. Locating the initial stages of speech-sound processing in human temporal cortex. *NeuroImage* 2006; **31**:1284–1296.
22. Banai K, Nicol T, Zecker SG, Kraus N. Brainstem timing: implications for cortical processing and literacy. *J Neurosci* 2005; **25**:9850–9857.
23. Wong PC, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 2007; **10**:420–422.
24. Suga N, Ma X, Gao E, Sakai M, Chowdhury SA. Descending system and plasticity for auditory signal processing: neuroethological data for speech scientists. *Speech Commun* 2003; **41**:189–200.
25. Xu Y. Contextual tonal variations in Mandarin. *J Phon* 1997; **25**:61–83.