Perception of Pitch Contours by Native and Non-Native Tone Listeners

Ratree Wayland, Yiqing Zhu and Edith Kaan
Department of Linguistics, University of Florida, USA

ABSTRACT

Level, rising and falling pitch contours were presented to twenty native speakers of Mandarin Chinese and twenty-one native speakers of English for discrimination in a same-different categorial discrimination task. Overall, Mandarin listeners were significantly more successful than English listeners. However, both groups exhibited a similar pattern of perceptual advantage for rising and falling contours. They were also more successful at discriminating the level contour from the rising contour than from the falling contour. Experience with the native tone systems may partially explain the results. However, a rising contour may be inherently more perceptually salient than either a falling or a level contour.

Keywords: Pitch contour perception, Mandarin, English

1. INTRODUCTION

Native and nonnative listeners of a tone language rely on different pitch dimensions when perceiving tones. It has been shown, for example, that native listeners of Thai and Yoruba place more weight on pitch direction and contour whereas native English listeners put more emphasis on average pitch and extreme end points [1]. A later study [2] further confirmed native English listeners' greater reliance on average pitch over pitch contour in tone perception. Results of recent ERP studies are also consistent with these earlier behavioral findings [3, 4]. However, relative degree of discriminability of different pitch contour types among native tone listeners remains to be explored.

The aim of this study was to deepen our understanding of pitch contour perception among native and non-native tone listeners. Specifically, we aim to compare and contrast the ability to discriminate rising, falling and level pitch contours among native Mandarin and native English listeners. Even though native listeners of a tone language are better at discriminating lexical tones of a foreign language than non-native tone listeners [5], their discrimination of non-linguistic pitch (pure tone) stimuli with a linearly falling or rising contour may be worse than that of non-native tone listeners [6]. In this study, we tested native Mandarin and native English listeners on their ability to discriminate linguistic pitch stimuli. These stimuli were neither natural lexical tone stimuli nor were they non-linguistic in nature. They were nonsense CV (/ba:/) syllables with linearly falling, rising or level pitch contours. The two main questions that guided the research were:

1. Would discrimination of each linguistic pitch contour be comparable among Mandarin and English listeners?
2. Would both groups of listeners show comparable or differential degrees of discrimination for all pitch contours?

2. METHODS

2.1. Stimulus Materials

A /ba:/ syllable produced by a male, native speaker of Assamese was acoustically modified using Praat software [7]. The fundamental frequency (F_0) of this original /ba:/ syllable was manipulated to generate 17 tokens with a linearly rising, 17 tokens with linearly falling and 17 tokens with level F_0 contour (Table 1-3). F_0 onset and offset difference (i.e., the slope) was 15 Hz for the falling and rising contour types. All stimuli were 250ms long with a 20ms amplitude on- and off-ramp to remove any existing transient noise. They were sampled at 44,100 Hz and normalized for peak intensity (98% of the full scale) before presentation.

Table 1: F_0 onset and offset values (in Hz) of stimuli with falling contour used in the study

Stimulus no.	F0 onset	F0 offset	Contour
ba:1	111	96	falling
ba:2	115	100	falling
ba:3	119	104	falling
ba:4	123	108	falling
ba:5	127	112	falling
ba:6	131	116	falling
ba:7	135	120	falling
ba:8	139	124	falling
ba:9	143	128	falling
ba:10	147	132	falling
ba:11	151	136	falling
ba:12	155	140	falling
ba:13	159	144	falling
ba:14	163	150	falling
ba:15	167	154	falling
ba:16	171	156	falling

ba:17	175	160	falling

Table 2: F_0 onset and offset values (in Hz) of stimuli with rising contour used in the study

Stimulus no.	F0 onset	F0 offset	Contour
ba:18	96	111	rising
ba:19	100	115	rising
ba:20	104	119	rising
ba:21	108	123	rising
ba:22	112	127	rising
ba:23	116	131	rising
ba:24	120	135	rising
ba:25	124	139	rising
ba:26	128	143	rising
ba:27	132	147	rising
ba:28	136	151	rising
ba:29	140	155	rising
ba:30	144	159	rising
ba:31	150	163	rising
ba:32	154	167	rising
ba:33	156	171	rising
ba:34	160	175	rising

Table 3: F_0 onset and offset values (in Hz) of stimuli with level contour used in the study

Stimulus no.	F0 onset	F0 offset	Contour
ba:35	96	96	level
ba:36	100	100	level
ba:37	104	104	level
ba:38	108	108	level
ba:39	112	112	level
ba:40	116	116	level
ba:41	120	120	level
ba:42	124	124	level
ba:43	128	128	level
ba:44	132	132	level
ba:45	136	136	level
ba:46	140	140	level
ba:47	144	144	level
ba:48	150	150	level
ba:49	154	154	level
ba:50	156	156	level
ba:51	160	160	level

2.2. Participants

Twenty (10 male, 10 female) native Mandarin and twenty one (3 male, 18 female) native English listeners participated in the study. All participants were between ages of 19-40. All passed a hearing screen at octave frequency from 250-4000 Hz as administered on site. None reported any speech or neurological impairment.

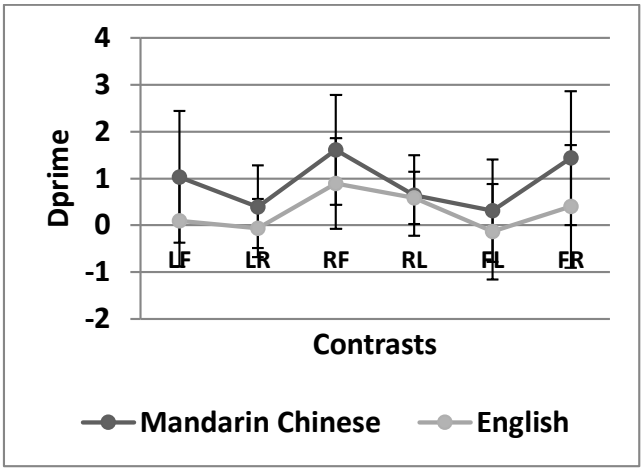
2.3. Procedures

Participants' discrimination of all possible pair-wise comparisons, yielding six contrasts in total (level-fall, level-rising, rising-falling, rising-level, falling-level, falling-rising), among the three pitch contours was examined in a same-different categorial discrimination task. For this task, participants heard four stimuli per trial with a 500ms inter-stimulus interval. Two types of trials were used; different trials and same trials. In different trials, the F_0 contour (falling, rising, or level) of the first three stimuli was the same, but that of the fourth stimulus was different from that of the first three (e.g., ba:1, ba:3, ba:5, ba: 18). In same trials, on the other hand, the F_0 contour of all three stimuli were the same (i.e., ba:1, ba:3, ba:5, ba:7). The participant's task was to decide whether the pitch contour of the last stimulus was the same or different from that of the first three. A 'different' response counted as a hit for 'different' trials, but as a 'false alarm' for same trials. D prime scores [Z (hit rate) - Z (false alarm rate)] were computed for each participant for statistical analyses. A 1.0 hit rate and a 0.0 false alarm rate was corrected by the formula 1-1/2N and 1/2N respectively, where N equals the number of trials.

3. RESULTS

Dprime scores obtained from both Mandarin and English listeners are displayed in Figure 1. Level-falling (LF) dprime score reflected participants' ability to discriminate level contour from falling contour (i.e., when the pitch contour of the last stimulus was falling, but that of the first three stimuli was level) while FL dprime score reflected their ability to discriminate the last stimulus with a level contour from the falling contour of the preceding three stimuli, etc. As evident in this figure, Mandarin listeners performed better than English listeners. In addition, both groups show similar pattern of discrimination sensitivity with RF being the easiest and FL being the most difficult contrast respectively.

Figure 1 Discrimination D prime scores and SD for all 6 pitch contour contrasts (level-falling, level-rising, rising-falling, rising-level, falling-level and falling-rising) for both Mandarin and English listeners.



A repeated measures ANOVA with native Language (L1) (2 levels; Mandarin and English) as the between-subject factor and pitch contour Contrast (6 levels; LF, LR, RF, RL, FL, FR) as the within-subject factor yielded significant main effects of Language [F (1, 39) = 11.68, p < .001], and of pitch contour Contrast [F (5, 195) = 8.43, p <.001]. The Language x Contrast interaction was, however, non-significant [F (5, 195) = 1.40, p=.23].

Table 4 below shows mean dprime values averaged across both groups of listeners. Post-hoc pair-wise comparison showed that d prime score for RF was significantly (p<.05, Bonferroni adjusted) higher that those of other contrasts except FR suggesting that discrimination between rising and falling pitch contours was easiest. In addition, d prime scores for RL and FR were significantly higher than that of FL indicating that the listeners find it easier to differentiate a rising contour from a level contour than to discriminate between a falling and a level pitch contour.

Table 4: Mean Dprime values and standard errors for all 6 pitch contour contrasts across both groups of listeners

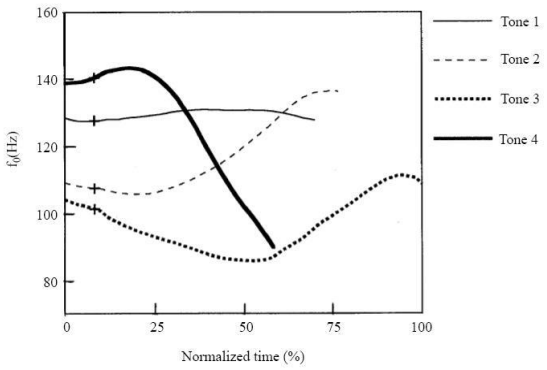
Pitch Contour	D prime	Standard Error
Level-falling (LF)	.565	.189
Level-rising (LR)	.167	.118
Rising-falling (RF)	1.253	.167
Rising-level (RL)	.610	.112
Falling-level (FL)	.088	.165
Falling-rising (FR)	.920	.214

To further explore relative degree of perceptual saliency between rising and falling pitch contours, a second repeated measures ANOVA comparing dprime scores of LR and RL contrasts against those of LF and FL contrasts was performed with Language as the between subject factor. The results obtained revealed that Mandarin listeners were significantly better than the English listeners [F(1, 80) = 7.52 p <.01], but that their discrimination of both rising and falling pitch contours (from a level pitch contour) was comparable [F(1, 80) = .23, p= .64] and no significant interaction between Language and tone Contrast was found [F(1, 80) = 2.77, p =.10].

4. DISCUSSION

The results obtained suggested that overall, native Mandarin listeners were superior to native English listeners in their ability to categorially discriminate pitch contours. Experience with native Madarin tone system may in part explain this result. It is interesting, however, that both groups of listeners found a rising-falling and a falling-rising pitch contour contrasts the easiest to discriminate. Specifically, they were better able to differentiate a falling contour from previously heard rising contours and vice versa. As shown in Figure 2 below, the presence of Mandarin Tone 2 (rising), Tone 3 (falling rising) and Tone 4 (falling) may have sensitized Mandarin listeners to the contrast between the rising and the falling pitch contours.

Figure 2: Mandarin Chinese tone (from Xu, 1997)



Native English listeners, on the other hand, may have benefited from the presence of rising and falling intonations used in English to differentiate a question

from a statement, for example. However, this explanation should be taken with caution since intonations typically extend over a linguistic unit longer than 250ms.

It is also possible that falling and particularly rising pitch contours are psychoacoustically more salient than a level pitch contour since both groups of listeners, despite their different L1 background, find them to be easily discriminable.

It is also interesting that both groups of listeners found it easier to discriminate a level pitch contour from a rising pitch contour than from a falling pitch contour, suggesting that a rising pitch contour may be more salient than a falling pitch contour.

Evidence exists to support the finding that a rising pitch contour may be more perceptually salient than a falling pitch contour. Brain stem's frequency-following-response (FFR) was found to be more faithful and robust for rising pitch contour (Mandarin Tones 2 and 3) than for level and falling pitch contours (Mandarin Tone 1 and 4) [10]. The FFR amplitude for rising tonal sweeps is larger than that of falling tonal sweeps [11].

However, the finding that only RL, but not LR was more discriminable than FL suggests an order of presentation effect. Specifically, a rising pitch contour is more discriminable when it precedes rather than follows a level pitch contour. This order of presentation effect has also been reported in previous study [12] in which it was found that Cantonese listeners significantly more sensitive to pitch difference between two syllables when the pitch of the first syllable was higher than that of the second.

This order of presentation effect may explain a lack of a significant difference in d prime scores between RL and LR contrasts versus FL and LF contrasts we reported.

In sum, both language-specific and language-general factors may explain the patterns of results obtained. Due to native language experience, Mandarin listeners outperformed native English listeners in their overall performance. However, due perhaps to relative degrees of psychoacoustic saliency across pitch contour types, both groups of listeners exhibited similar patterns of discrimination. A rising pitch contour may be more perceptually salient than a falling and a level pitch contour. However, the presence of a presentation order effect warrants further experimentation.

5. REFERENECES

[1] Gandour, J., Harshman, R. 1978. Cross-language difference in tone perception: A multidimensional scaling investigation. *Language and Speech* 21, 1-33.

[2] Gandour, J. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics 11, 49-75*.

[3] Kaan, E., Wayland, R., Bao, M., Barkely, C. 2007. Effects of native language and training on lexical tone perception: An ERP study. *Brain Research* 1148, 113-122.

[4] Chandrasekaran, B., Gandour, J., Krishnan, A. 2007. Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling of the mismatch negativity. *Restorative Neurology and Neuroscience* 25, 195-210.

[5] Wayland, R., Guion, S. 2004. Training native English and native Chinese speakers to perceive Thai tone. *Language Learning* 54, 681-712.

[6] Bent, T., Bradlow, A. 2006. The influence of linguistic experience on the cognitive processing of pitch in speech and non-speech sounds. *Journal of Experimental Psychology: Human perception and performance* 32(1), 97-103.

[7] Boersma, P., Weenink, D. 2008. *Praat (5.0.4).*Doing phonetics by computer. Retrieved from http://www.fon.hum.uva.nl/praat/ on March 21, 2008

[8] Abramson, A.S. 1962. *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments*. Bloomington: Indiana U. Res. Center in Anthropology, Folklore, and Linguistics, Pub. 20.

[9] Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.

[10] Krisnan, A., Xu, Y., Gandour, J.T., Cariani, P. 2004. Human Frequency following response: representation of pitch contours in Chinese tones. *Hearing Research*, 189,1-12.

[11] Krishnan, A., Parkinson, J., 2000. Human frequency-following responses: representation of tonal sweeps. *Journal of Audiology & Neuro-Otology* 5, 312-321.

[12] Francis, A., Ciocca, W. 2003. Stimulus presentation order and the perception of lexical tones in Cantonese. *J. Acoust. Soc. Am.* 114(3), 1611-1622.

[1] We would like to thank Dean Mercado, Kayla Johnson and Lauren Concepcion for their assistance with the data collection.