EFFECTS OF NATIVE LANGUAGE ON PERCEPTION OF LEVEL AND FALLING TONES

Bin Li & Lan Shuai

Department of Chinese, Translation and Linguistics, City University of Hong Kong, Hong Kong
binli2@cityu.edu.hk

ABSTRACT

In this paper, we compare tone perception by speakers from different language backgrounds (L1s) including Mandarin, Cantonese, Thai, and Indo-European languages such as English, French, and German. We focus on how two phonetic cues i.e., pitch height and pitch slope, affect perception of pitch direction. Significant differences are found in the degrees of sensitivity to the two cues across L1 groups.

Keywords: tone, speech perception, pitch direction, pitch height, pitch slope

1. INTRODUCTION

Native languages affect their speakers' perception of the world [3], as well as of sounds from other languages such as tones [6]. However, there is no consensus yet on the mechanism of how L1 functions in non-native tone perception. Models have been developed and applied to address this issue [1, 3], for example, the Attention to Dimension Model (A2D) and the Perceptual Assimilation Model (PAM). A2D focuses on degrees of attention to phonetic cues in speech perception, such as pitch height in tone perception, and PAM on how contrasts of phonological features affect the mapping of a non-native sound, e.g. tones, into the L1 categories. A2D explores at the phonetic level to explain the learning of a new phonetic category, but fails to capture the effect resulted from the speaker's phonological knowledge of their L1. PAM, on the other hand, addresses this latter issue, but lacks the power to explain how individual non-native phonemes are identified and perceived. In recent development of both models, authors in [4] and [11] found that both the phonological and phonetic cues are important in perceiving non-native tones. Admitting the crucial roles of cues at both levels, we are more interested in how they functions jointly in tone perception. As a first step, we try to find out how such cues are processed in the judgment of pitch direction by speakers from various language backgrounds.

2. EXPEREMENTAL DESIGN

A recent study [9] found that Mandarin speakers pay more attention to pitch change than Cantonese speakers when identifying synthesized tones. The current work is a replica of this previous study with speakers from various language backgrounds to study roles of L1. Our study concerns two Mandarin tonal categories, the high level tone (T1) and the high falling tone (T4). A series of pitches are synthesized by changing the values of pitch height and pitch slope of a base syllable. Native speakers from four language groups are recruited to complete an ABX identification task.

We use 4-way mixed design ANOVA in analyzing results. Phonetic cues of onset F0 and pitch slope (△F0) are used as within-subject factors. Another within-subject factor is the inter-stimulus interval (ISI) with two levels, 500ms and 1500ms. L1 is used as the between-subject factor with four levels.

2.1. Stimulus materials

We record two meaningful monosyllabic Mandarin words bā 八 "eight" (T1) and bà 霸 "a bully" (T4) produced by a male native speaker of Mandarin. Each syllable is repeated nine times. One clear utterance of bā 八 "eight" (T1) is chosen as the base for pitch synthesis. The base syllable is modified using Praat [2]. First, its duration is adjusted to 500ms. Secondly, its pitch value is modified along two dimensions, F0 and △F0, at 0ms, 100ms and 500ms. F0 refers to the average pitch of the first 100ms from onset. △F0 stands for pitch difference between 101ms and 500ms. Thirdly, a 500Hz low-pass filter is applied to remove interference of segmental information.

The preparation of the stimulus materials replicates those used in [9]. Thirty six stimuli are synthesized: 6 steps in F0 synthesis from 100Hz to 130Hz with a step of 6Hz and 6 steps in △F0 synthesis from 0Hz to 30Hz with a step of 6Hz (In mel scale, 6 Hz difference between 70 and 130 Hz equals to an average of 8.46 mel (S.D. 0.19 mel)). For example, when F0=100Hz, with △F0 of 0Hz, 6Hz, 12Hz, 18Hz, 24Hz, and 30Hz, we have 6 pitches that all start at 100Hz and end at 100Hz, 94Hz, 88Hz, 82Hz, 76Hz, 70Hz respectively.

2.2. Participants

Eighty seven participants take part in the experiment. They all fill out a language background questionnaire before testing and information summary is listed in Table 1.

Table 1: Background information of all participants.

L1	#	Age (S.D.)	M/F	Yrs. of learning Mandarin
Mandarin	29	24.7(3.3)	13/16	n/a
Cantonese	28	21.6(2.2)	15/13	9.2(5.5)
Thai	9	23.2(2.0)	5/4	2.9(1.8)
Indo-European	21	23.1(7.2)	12/9	1.2(1.2)

2.3. Tasks

ABX identification paradigm is used. Each test trial contains a sound triplet: A and B are recordings of T1 or T4 respectively, as two references; X is the target sound that participants need to compare to A/B and identify.

The two reference sounds are chosen from the representative repetition of bā (T1) and bà (T4) by the same speaker. They are treated as a prototype of the level and falling tonal categories. The onset pitch of T1 falls in the range of onset F0 of our stimuli, between 100Hz and 130Hz; and the range of pitch slopes of T4 covers the full range of our stimuli, between 70Hz and 130Hz. Both reference sounds are normalized to 500ms and go through a 500Hz low-pass filter. The orders of T1/T4 presentation are counterbalanced for each participant. The X sound is drawn from the pool of 36 stimuli and presented randomly.

Participants are asked to judge whether X is similar to the A or B. Then they are asked to do a comparison between the target and the reference sound they select, on a scale of 5 with 1 being "very similar". There are 10 practice trials in the beginning of the test. Each participant completes two sessions, ISI 500ms and ISI 1500ms. Each session contains 36 trials. The sequence of the two conditions is counterbalanced across participants.

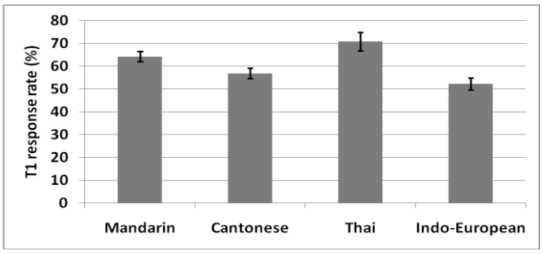
3. RESULTS AND DISCUSSION

We adopt 4-way repeated-measures ANOVA to both the level/falling (T1/T4) response and the similarity rating, with F0, △F0, ISI as the within-subject factors, and the language background as the between-subject factor. The T1/T4 response is transformed to the percentage of the T1 responses. The raw similarity rating is also transformed to the percentage of the T1 responses, as the similarity rating is a more detailed response to the T1/T4 identification. Due to the space limitation, we only discuss the T1/T4 response results here.

3.1. Results

First of all, results show a main effect of L1 on the percentage of T1 responses (F (3, 83)=7.668, p<0.001, η_p^2=0.217). This suggests significant differences across language groups in their perceptual patterns of the synthesized pitches. The mean scores of all four groups are shown in Figure 1. On average, Thai and Mandarin speakers are more likely to label the synthesized pitches as level. A post-hoc comparison reveals a significant difference between Thai and Cantonese speakers (p<0.027), and between Thai and Indo-European language speakers (p<0.002). The responses of Mandarin speakers also differ significantly from the Indo-European language speakers (p< 0.007).

Figure 1: The mean percentage of T1 responses by four groups of language speakers.



Secondly, main effects are also found for both F0 (F (2.421, 200.962)=9.023, p<0.001, η_p^2=0.098) and △F0 (F(2.385, 197.962)=77.529, p<0.001, η_p^2=0.483) on percentage of T1 responses. There is also a significant interaction between the two factors (F(18.131, 1504.879)=2.093, p<0.005, η_p^2=0.025). The mean percentages of T1 responses for all 36 stimuli by each language group are shown in Figure 2. Thirdly, statistics reveal a significant interaction between L1 and △F0 (F(2.385, 197.962)=3.248, p<0.003, η_p^2=0.105); but no significant interaction between L1 and F0 (F(7.264, 200.962)=1.117, p=0.354, η_p^2 =0.039).

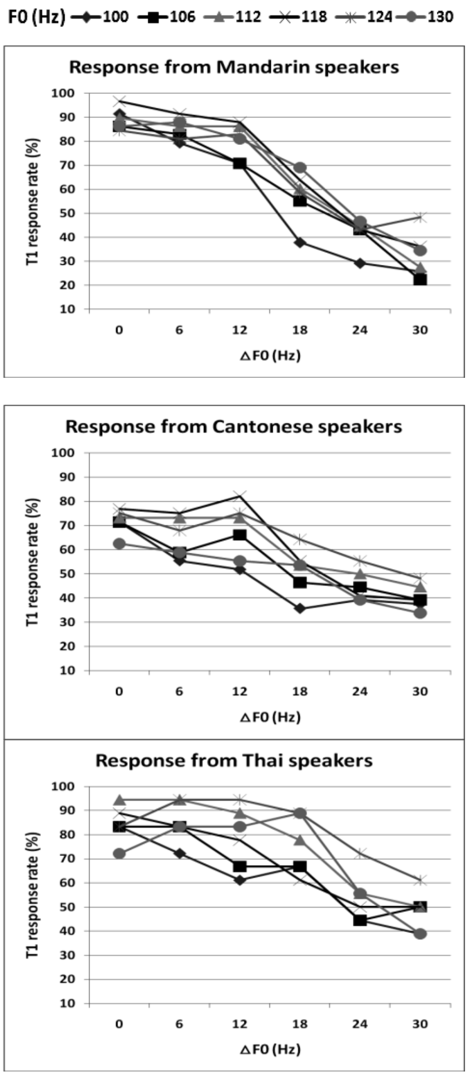
Last but not least, there is a significant interaction between ISI conditions and F0 (F(2.752, 228.387)=8.959, p<0.001, η_p^2=0.097). Since the ISI is not the focus of the current study, we will not discuss this factor in further detail. The values in Figure 2 are averaged across the two ISI conditions.

3.2. Discussion

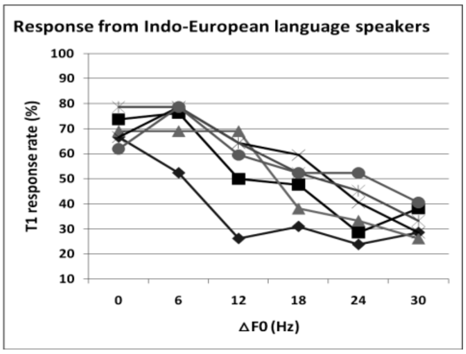
Generally speaking, all groups' performance is proportionately correlated to △F0. The larger the △F0 is, the less likely that a token is labeled as being the level tone (T1). That is to say, the greater the pitch slope is, the more likely a token is perceived as the falling tone (T4). Specific perceptual patterns vary greatly across language groups.

Figure 2: Mean percentage of the T1 responses for all stimuli in each group of language speakers.

(a) (b) (c)



(d)



Perception by native speakers of Mandarin manifests a near-categorical pattern. Result lines indicating 6 onsets in Figure 2(a) are closely clustered to each other, suggesting little effects on perception from F0 onsets. On the other hand, performance seems more affected by △F0. When △F0 is greater than 24Hz, tokens are more likely to be labeled as falling; whereas those smaller than 12Hz as level. When △F0 is 18Hz, their judgment is mixed. This pattern may be explained from the fact that Mandarin has one level tone and three contour tones. Pitch direction is the cue in classifying tones while pitch onsets are not functionally contrastive. Mandarin speakers therefore mainly rely on degrees of pitch change to distinguish level from falling tones.

Cantonese speakers perform in a similar way in general but seem more sensitive to F0 cues than Mandarin speakers. For Cantonese speakers, F0 seems not determining in tone differentiation but does exert a somewhat stronger effect than in the Mandarin group, as the lines in Figure 2(b) are not as tightly grouped as in Figure 2(a). Cantonese speakers are also sensitive to △F0. A pitch change of 12Hz, or smaller, does not affect tone labeling; but a change greater than 18Hz results in a change in tone classification. Moreover, Cantonese speakers may not be as confident as Mandarin speakers in their judgments because percentages of the former group in level/falling responses fall in a range that is no greater than 50%.

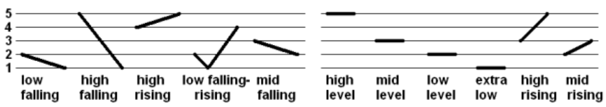
Thai speakers, as shown in Figure 2(c), exhibit a categorical-like trend as Mandarin speakers, but with a different level/falling "boundary". They are more likely to label stimuli with a △F0 that are smaller than 18Hz as level, and those with a △F0 greater than 24Hz as falling. This pattern is distinct from all other groups. It seems that Thai speakers are not significantly affected by △F0 as long as it does not exceed 18Hz, which becomes a boundary between level and falling pitches. Moreover, F0 onsets also exert influences on their percentage of level/falling responses but not the general categorization of level and falling.

Similarly as tonal language speakers, Indo-European language speakers as shown in Figure 2(d), tend to label tokens with greater △F0 as falling. In addition to this, they seem more sensitive to F0 onset, as they tend to label stimuli with the lowest F0 onset as falling regardless various pitch slopes. Response lines of all onsets are closer to each other, but more separated than Mandarin speakers. This suggests some degree of constant dependence on both phonetic cues; and a sharper sensitivity to pitch height than that of Mandarin speakers.

Except for Mandarin speakers whose maximum and minimum percentages in responses range from 20% to over 90%, the max-min range of other groups is no greater than 55%. Cantonese speakers' judgment covers the smallest range of less than 50% and with less-skewed lines. This pattern may suggest lack of confidence in judgment or a lower sensitivity to △F0. Meanwhile, Thai speakers' categorical-like, yet wide-spread, judgment suggests functions from both parameters.

Patterns found in Cantonese and Thai may be interpreted as a result from experience in their native tonal systems. Thai and Cantonese tones are shown in a 5-scale system in Figure 3 [7, 8, 10]. Cantonese contains four level tones and therefore its speakers mainly rely on F0 for distinction. For Thai, its high, mid, and low tones differ in pitch direction. Therefore, its speakers need both pitch onset and slope to distinguish tones. This may help explain why our Thai speakers are sensitive to F0, and to △F0 at a certain range in perceiving pitch direction.

Figure 3: Tones in Thai (left) and Cantonese (right).



4. CONCLUSION

This study examines how speakers of various L1s distinguish pitch direction in order to study L1's effects on tone perception. It is found that L1 influences sensitivity to onset pitch height and pitch slope in identifying level/falling pitches.

Our language speakers show clear and varied degrees of dependence on pitch slope. In general, they all rely on the phonetic cue of pitch change to make judgment on tones. Their boundaries between level and falling tones also differ from each other. On the other hand, F0 onsets' influence is language specific.

5. ACKNOWLEDGEMENTS

This study is supported by CLASS grant #9610158 from the City University of Hong Kong. We also thank Dr. Liang Lei for his assistance in data collection.

6. REFERENCES

[1] Best, C.T. 1995. A direct realistic view of cross-language speech perception. In Strange, W. (eds.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research.* Baltimore: York Press, 171-204.

[2] Boersma, P., Weenink, D. 2011. Praat: Doing phonetics by computer [Computer program], version 5.2.16. *http://www.praat.org*

[3] Carroll, J.B. **(**eds.**)** 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf.* MIT Press, Cambridge, MA.

[4] Francis, A.L., Ciocca, V., Ma, L., Fenn, K. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36, 268-294.

[5] Francis, A.L., Nusbaum, H.C. 2002. Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 28, 349-366.

[6] Gandour, J. 1983. Tone perception in far Eastern languages. *Journal of Phonetics* 11, 149-175.

[7] Ladefoged, P. 2001. *Vowels and Consonants: An Introduction to the Sounds of Languages.* Malden, Mass.: Blackwell.

[8] Ladefoged, P., Johnson, K. 2011. *A Course in Phonetics.* Boston, MA: Wadsworth Cengage Learning.

[9] Li, B., Zhang, C. 2010. Effects of F0 dimensions in perception of Mandarin tones. *Proceeding of ISCSLP* Tainan.

[10] Matthews, S., Yip, V. 1994. *Cantonese: A Comprehensive Grammar.* London: Routledge.

[11] So, C.K., Best, C.T. 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53, 273-293.