THE INFLUENCE OF LINGUISTIC BACKGROUND ON ASYMMETRIC TONE PERCEPTION INDUCED BY POSITION OF CONTEXT

Hong-Ying Zheng[a,b], *Ching-Pong Au*[c] *& William S-Y. Wang*[a]

[a]Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong;
[b]Department of Software Engineering, Shenzhen Institute of Information Technology, Hong Kong;
[c]Division of Language Studies, Community College of City University, Hong Kong
hommyzhy@gmail.com; chingpau@cityu.edu.hk; wsywang@ee.cuhk.edu.hk

ABSTRACT

Previous studies on categorical perception of lexical tones showed that perception of Cantonese lexical level tones depends on pitch level of contexts and the contextual effect is specific to speech. Moreover, a pair of stimuli with a right context is easier to discriminate than that with a left context. Although this directional asymmetric perception was observed from both native and nonnative tone language speakers, it is not clear yet whether such effect is also present in speakers without any tone experience.

This paper investigates this question further by comparing the perception between native Cantonese speakers and French speakers on Cantonese level tones, which are presented in left and right contexts. The results suggest that the asymmetric effect is due to both language experience and general auditory processing. The second aim of this paper is to examine whether categorical perception of lexical tones also depends on lexical meaning of the syllable, by comparing the performance from native speakers on real speech (with valid lexical meaning) and pseudowords (without valid lexical meaning).

Keywords: tone, categorical perception, position of context

1. INTRODUCTION

The acoustic features of speech sounds vary significantly across different contexts and speakers, yet listeners typically perceive speech without difficulty. Categorical perception (CP) of speech occurs when listeners map these acoustic signals to discrete categories [3], resulting in better discrimination of stimuli across category boundaries than equivalently separated stimuli within the same category.

In tone languages, words are made up tones (tonemes), as well as consonants and vowels (phonemes). The most representative features of lexical tones are pitch contours. For example, different levels of pitch height in Cantonese determined three level tone categories: high-level (tone[1], /55/), mid-level (tone[3], /33/) and low-level (tone[6], /22/) [1].

Perception of a phoneme typically depends on features of the target itself. However, when a specific signal is embedded in different types of contexts, the perception is different. For example, a same high-level tone will be perceived as a high level tone, a high rising tone, or a falling rising tone depending on the pitch heights of its following neutral tones [4].

Although there are quite a few studies focousing on tone CP, most of them are on citation form [2, 6, 7].Only two studies [8, 9] consider the contextual effect on tone CP. In these two studies, contexts from different positions (proceeding the target, the left context sentence: LC; or following the target, the right context sentence: RC) have different effects on both the degree of CP and discriminability by showing better discrimination from RC continua than from LC continua. Moreover, stronger CP is observed from native subjects than that from speakers of another tone language. Since both groups of subjects have tone experience, it remains unknown whether the effects of the contexts regarding different positions are only influenced by linguistic experience. Moreover, even though earlier studies indicate that CP of lexical tones is present in real speech but not in nonspeech, it is not clear whether the lexical meaning of the syllable plays a role in the CP.

Two experiments were carried out, following up the experiments in [8] and [9] by comparing the stimuli embedded in LC and RC sentences. In Experiment I, two different groups of subjects—tone language (Cantonese) speakers and non-tone language (French) speakers—perceived both speech and nonspeech stimuli. In Experiment II, native Cantonese speakers were recruited to perceive pseudoword stimuli. Two experiments only differ in target and subjects.

2. EXPERIMENT I AND II

2.1. Subjects

Sixteen university students from two different language backgrounds, with no reported history of speaking or hearing disability, participated in the experiment. Eight were native Cantonese speakers, (4 F, aged: 22±2.2). Eight were French speakers (4 F, aged: 21±2.3). All subjects were paid to participate. French speakers have no experience on Cantonese. French speakers only attended in Experiment I, while Cantonese speakers attended in both Experiment I and II.

2.2. Stimuli

In Experiment I, four sets of stimuli with a 2 Targets [TS] (speech and nonspeech) $\times 2$ Contexts (left context [LC] and right context [RC]) design were tested. In Experiment II, two sets of stimuli with pseudoword as Targets in the same LC and RC contexts as those used in Experiment I were tested. The real speech stimuli in the present experiment were the same as those in early experiment [8]. LC sentence was /ni^1 go^3 zi^6 hai^6 TS/ (This word is TS) and RC sentence was /TS hai^6 mat^1 ji^3 si^1/ (What's the meaning of TS). The stimuli were all based on naturally uttered speech.

Each continuum contained 11 stimuli. The pitch contours of targets were fixed and heard as mid-level tones in the original speech. To make the targets were perceived as tones other than mid-level, the pitch contours of the contexts were varied at a step of 5 Hz upward or downward. After variation, the TS in the speech sentence may be heard as /fan^1/ ("to divide"), /fan^3/ ("to sleep") or /fan^6/ ("to share"), which differed from each other only in pitch value. Pitch contours were manipulated by PRAAT, the software for acoustic analysis (*http://www.fon.hum.uva.nl/praat*).

In each continuum, stimulus #7 was the baseline and the TS was expected to be perceived as a mid-level tone. TS in stimulus #1 and stimulus #11 were expected to be perceived as a high-level tone and a low-level tone respectively (refers to [8, 9]).

Synthetic nonspeech sound was made by concatenating the consonant portion from the syllable /*fan*/ and a saw wave. The pseudowords were synthesized by concatenating the consonant portion from the syllable /*fan*/ and the vowel portion from a naturally uttered syllable /*se*/ ("to write"). The duration of the TS from three types (speech, nonspeech and pseudowords) was the same. Moreover, their amplitude profiles were normalized. Therefore, except for the spectral information, these three sets of stimuli were kept as similar as possible.

2.3. Procedure

Stimuli were presented to listeners over a SONY headphone (MDR CD-777). Only discrimination tasks were carried out, because no corresponding phonemic categories are available for nonspeech or non-tone language speakers. The stimuli consisted of all pair-wise combinations of individual sentences separated by zero or two tokens along one continuum, with a 500ms Inter-Stimulus-Interval. There were a total of 29 such pairs for each continuum. Each pair repeated three times were distributed into 15 blocks randomly. Two pairs were separated by 6s Inter-Trial-Interval. The subjects were instructed to select 'yes' or 'no' on paper to indicate whether the TSs in a pair were the same or different.

2.4. Analysis

Three steps of analyses were carried out in the experiments. **(1)** To check whether there are peaks along the discrimination curves. A one-way repeated measures analysis of variance (ANOVA) with stimulus pair as the within-subject factor was carried out on each group per each stimulus set. **(2)**To examine whether there is a CP effect by comparing the discrimination performance between across-category pairs and within-category pairs. First, the discrimination scores were grouped into from across-category stimulus pairs (3,4,6,7) and from within-category pairs (1,2,5,8,9) based on the identification results from [9], which were also consistent with the obtained discrimination results shown in Fig. 1. After grouping, results from experiment I were shown in Fig. 2 and results from experiment II were shown in Fig. 3. A four-way mixed design repeated measures ANOVA with three within-subject factors (Target: speech vs. nonspeech; Position: LC vs. RC; Category: across-category vs. within-category) and one between-subject factor (Language: Cantonese vs. French) was carried out. **(3)** To examine the asymmetric effect regarding on context position on pseudowords by Cantonese speakers. In Experiment II, a two-way repeated measure ANOVA was carried out on the grouped discrimination performance, with Category and Position as two within-subject factors.

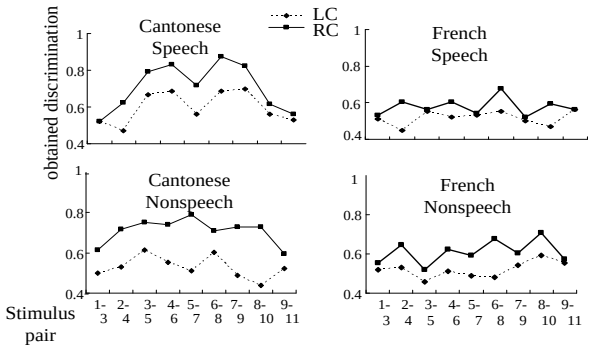
All post hoc analyses were conducted according to the general guidelines for analyzing effects in a three-factor design recommended in [5]. Greenhouse-Geisser correction was applied whenever it was necessary. Unless stated otherwise, all tests of significance were conducted at $p < 0.05$.

2.5. Results

2.5.1. Experiment I

The discrimination results were shown in Fig.1. The curves obtained from Cantonese speakers in both LC and RC continua were nonlinear, but only in speech condition [LC: F(8, 56)=4.302, p<0.05; RC: F(8, 56)=11.661, p<0.001]. No significant peaks along the discrimination curve in other continua indicated that only speech continua were perceived categorically by Cantonese speakers.

Figure 1: Discrimination of four sets of stimuli is averaged over each group of subjects.



Three two-way interaction effects [Position × Language: F(1,14)=4.642, p<0.05; Category × Language: F(1,14)=14.546, p<0.01; Target × Category: F(1,14)=13.868, p<0.01] on grouped results were obtained. The results showed that Cantonese and French discriminated differently on LC and RC continua. Moreover, the results also indicated that both language groups and the speech types influenced the discrimination difference between across-category pairs and within-category pairs. The post hoc analysis further revealed that although discrimination in RC was better than that in LC continuum for speakers from both language groups, the difference was more prominent for Cantonese speakers than to French speakers. Moreover, the CP effect was only present by Cantonese speakers but not by French speakers, and was present in speech condition but not in nonspeech condition.

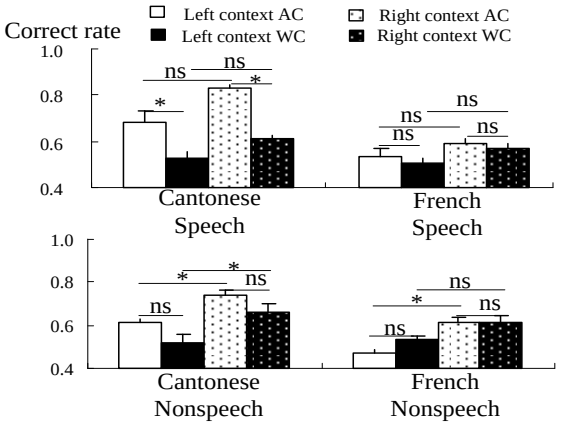
Two finer post hoc analyses using three-way repeated measures ANOVA with within-subject factors of Position, Category, and Target were carried out to examine the CP effect within each language group as shown in Fig.2. There was an interaction effect between Target and Category from Cantonese speakers [F(1, 7)=19.13, p<0.01]. Further analysis showed that speech was perceived categorically by Cantonese speakers but nonspeech was not. No Category effect was observed from French speakers, which indicated that French speakers perceived the stimuli non-categorically either in speech continua or in nonspeech continuum.

In summary, the comparison between the discrimination on speech and nonspeech continua from two groups of subjects revealed that:

(1) Discrimination is easier in RC than that in LC continua. Moreover, although this asymmetric discrimination performance regarding on the context positions is observed from both groups of subjects, the asymmetric effect is more prominent from Cantonese speakers than that from French speakers.

(2) CP effect is present from the performance by Cantonese speakers, especially in speech continuum. However no evidence shows that such effect is present by French speakers from any set of continuum.

Figure 2: Grouped discrimination according to across-category and within-category pairs by each group of subjects. "*" indicated statistically significant: p<0.05; "ns" indicated statistically nonsignificant.

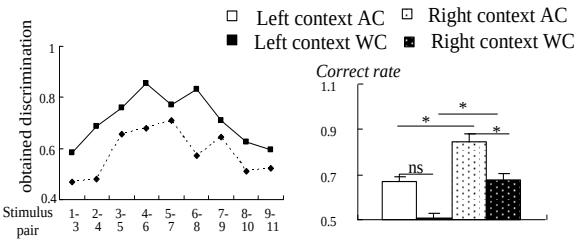


2.5.2. Experiment II

In Fig. 3, discrimination responses from LC and RC continua of pseudoword by Cantonese speakers were presented. The discrimination was nonlinear in both continua [LC: F(8, 56)=3.225, p<0.01; RC: F(8, 56)=4.44, p<0.01].

The statistical analysis on grouped discrimination showed two main effects [Position: F(1,7) =13.605 , p< 0.01; Category: F(1,7) = 43.074, p<0.001]. No interaction effect was obtained. These two main effects indicated that (1) pseudoword was perceived categorically by Cantonese speakers and (2) discrimination performance in RC was better than that in LC. This result was similar to that from real speech.

Figure 3: Discrimination curves and grouped discrimination performance according to across-category and within-category pairs from sets of pseudoword by Cantonese speakers. "*" indicated statistically significant: $p<0.05$; "ns" indicated statistically nonsignificant.



3. DISCUSSION

Experimental results show obvious peaks at the category boundaries from real speech and pseudoword continua by Cantonese speakers. The results further confirm that categorizing tones not only depends on the absolute pitch value of TS but also on the pitch distance between the TS and adjacent syllables, which is consistent with the findings on Mandarin tones [4].

There are no obvious peaks in nonspeech continua either by Cantonese speakers or by French speakers, or in speech continua by French speakers. However, speech and pesudoword continua have similar peaks by Cantonese speakers. The results suggest that the observed CP effect is due to language experience, which is mainly constrained to speech stimuli although not limited to valid lexical items. Without the linguistic experience, the same stimuli are not perceived categorically by non-tone language speakers. The CP effect due to linguistic experience by the native speakers does not generalized into the nonspeech continua, even though nonspeech targets share the same representative tone features, the pitch contours, with speech stimuli. Moreover, the results confirm our earlier findings [8, 9] that CP effect of the target in the contextual sentence is not due to the presence of a context, but due to the linguistic experience. Otherwise, the nonspeech stimuli should also be perceived categorically by native Cantonese speakers. Furthermore, the data suggests that CP of lexical tones is specific to speech stimuli which are valid phonemes in the language regardless the semantic meaning. Since subjects have the same degree of familiarity with pseudoword and nonspeech but nonetheless perceived differently, the familiarity of speech is not enough to explain the CP effect observed from both speech and pseudoword conditions.

The asymmetric perception due to the position of contexts observed in our earlier experiments [8, 9] is replicated. In the experiments, regardless linguistic background, discrimination performance from RC continua is generally better than that from LC continua. Taken into account of the same asymmetric pattern obtained from early study [9] even when the acoustic difference in RC and LC continuum was eliminated, we conclude that this position dependent asymmetric discrimination performance is relevant to both linguistic experience and general auditory processing.

4. ACKNOWLEDGEMENTS

Experiments were supported by Laboratoire Dynamique Du Langage, Lyon, France and Prof. P. W-M. Tsang, Department of Electronic Engineering, The City University of Hong Kong. We thank anonymous reviewers for comments.

5. REFERENCES

[1] Chao, Y.R. 1947. *Cantonese Primer*. Cambridge, MA: Harvard University Press.

[2] Francis, A.L., Ciocca, V., Ng, B.K.C. 2003. On the (non)categorical perception of lexical tones. *Perception & Psychophysics* 65(7), 1029-1044.

[3] Liberman, A.M., et al. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5), 358-368.

[4] Lin, T., Wang, W.S.Y. 1984. An experiment in tone perception (Trans.) (聲調感知問題). *Journal of Chinese Linguistics (中國語言學報)* 2, 59-69.

[5] Maxwell, S.E., Delaney, H.D. 2004. *Designing Experiments and Analyzing Data: A Model Comparison Perspective (*2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

[6] Wang, W.S.Y. 1976. Language change. In Harnad, S.R., Steklis, H.D., Lancaster, J. (eds.), *Origins and Evolution of Language and Speech*. New York: New York Academy of Sciences.

[7] Xu, Y.S., Gandour, J.T., Francis, A.L. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America* 120(2), 1063-1074.

[8] Zheng, H.-Y., et al. 2006. Perception of Cantonese level tones influenced by context position. *The 3rd International Conference on Speech Prosody Dresden*, Germany.

[9] Zheng, H.-Y., Tsang, P.W.-M., Wang, W.S.Y. 2007. Categorical perception of Cantonese tones in context: A cross-linguistic study. The *8th Annual Conference of the International Speech Communication Association (Interspeech 2007)* Antwerp, Belgium.