Cognitive Representation of Phonological Categories: The Evidence from Mandarin Speakers' Learning of Cantonese Tones

Kaile Zhang

Department of Chinese and Bilingual Studies,

The Hong Kong Polytechnic University

kaile.zhang@polyu.edu.hk

1. Background and the research question

(1) Speech production varies a lot. (see Figures 1 & 2)

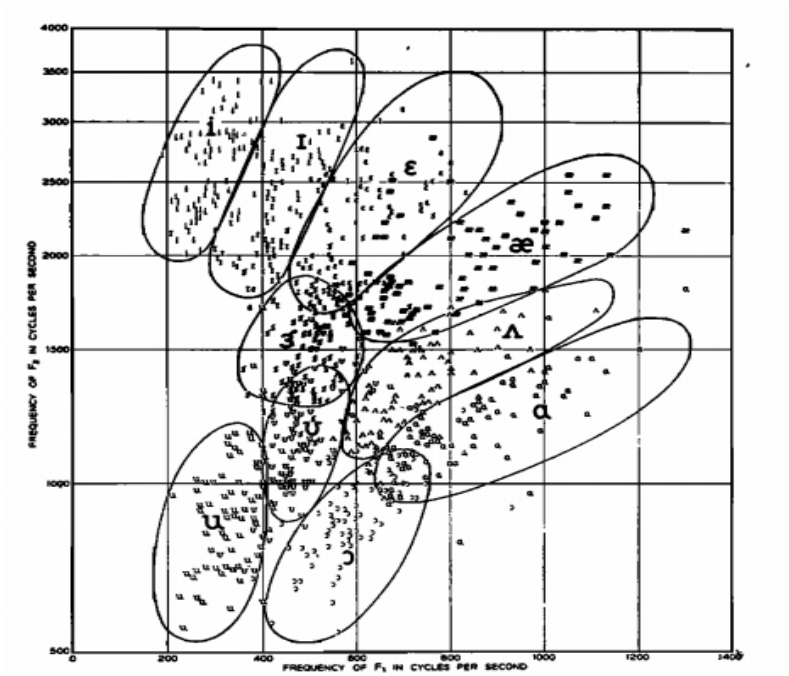


Figure 1. Frequency of second formant *versus* frequency of first formant for ten vowels by 76 speakers (Peterson and Barney, 1952).

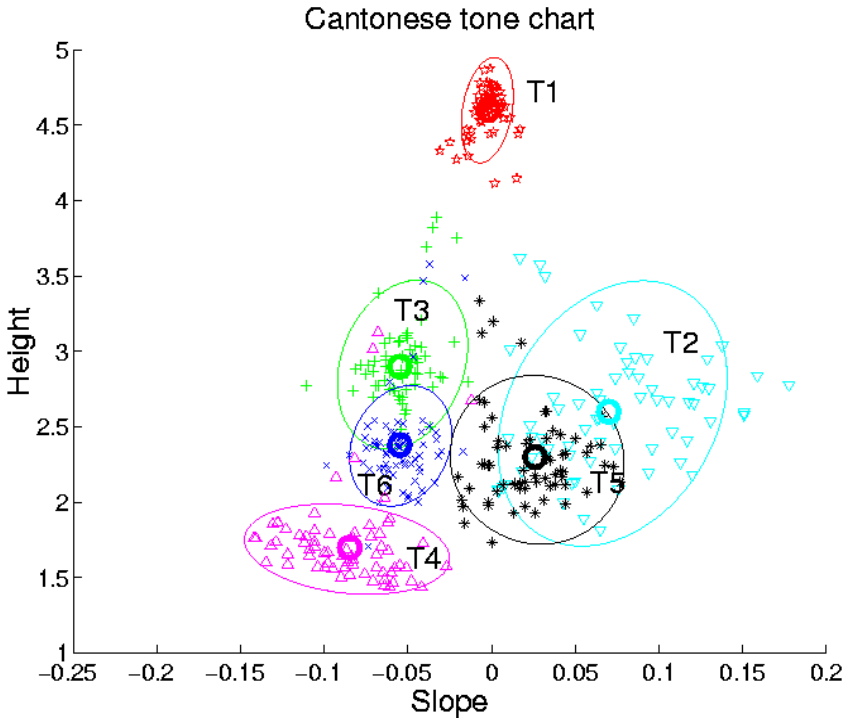


Figure 2. Production tone balloons for Cantonese (Peng, 2006).

(2) Speech perception.

The exemplar-based model (e.g., Goldinger, 1996; Johnson, 1997a; 1997b; Pierrehumbert, 2001)

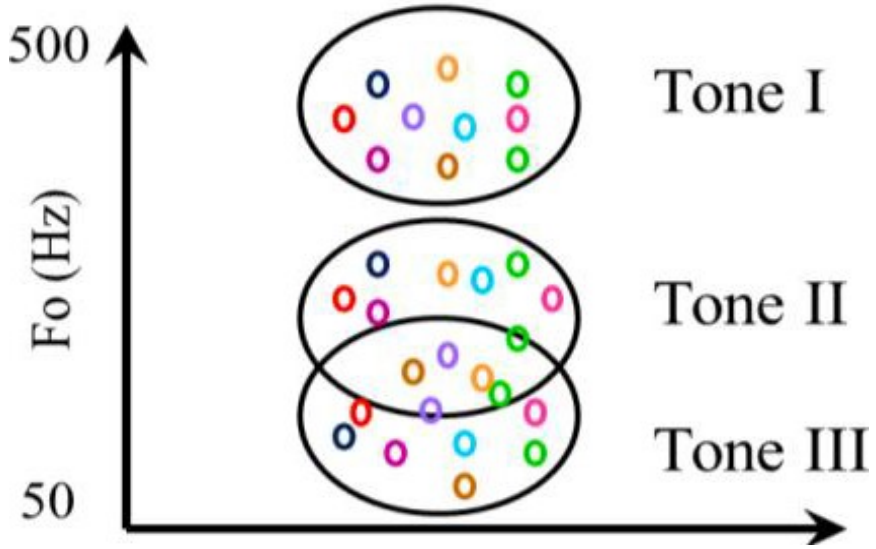
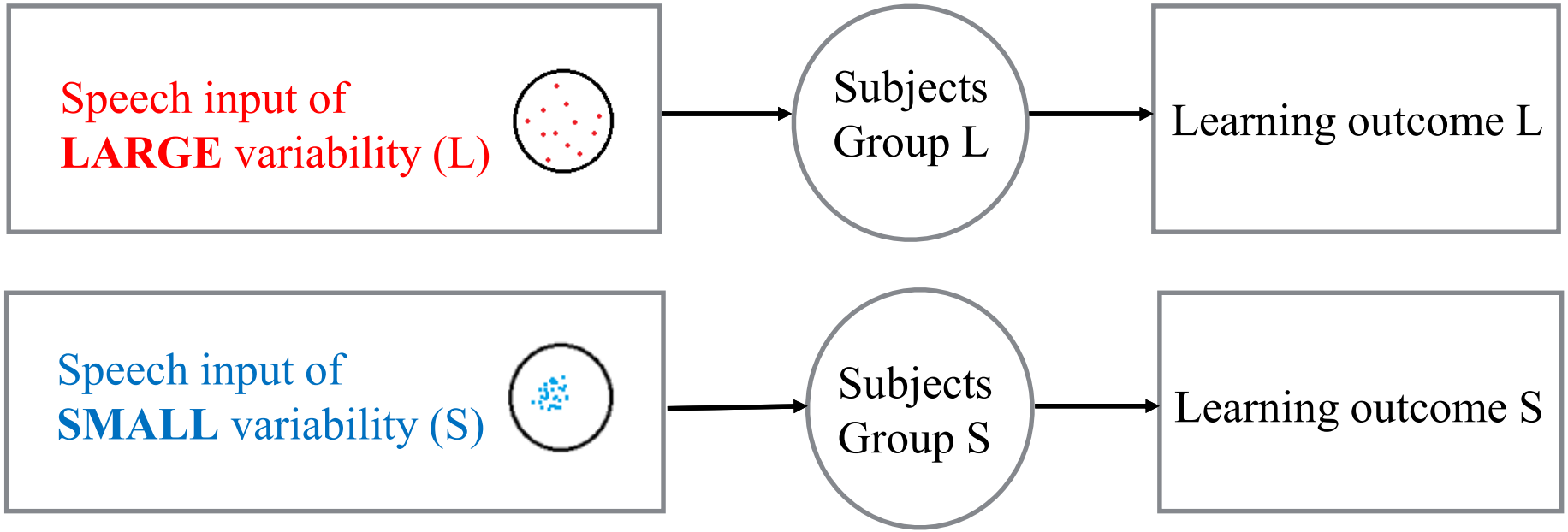


Figure 3. Summary of the exemplar-based model, exemplified for a system of three tones.

(3) Research question

What cognitive mechanisms are used to represent and perceive phonological categories.



2. Methods

2.1 Participants

(1) The experimental group

- 35 native Mandarin speakers (17F; North China; Undergraduate and graduate students; no musical training; with limited experience of Cantonese)
- 17 for Group L (large variability) and 18 for Group S (small variability)

(2) The control group

12 native Hong Kong Cantonese speakers (undergraduate students; no musical training)

2.2 Stimuli

(1) Recording:

Ten native Cantonese speakers (5F) × 36 syllables × 10 repetitions

		fan	fu	jan	ji	se	si
Tone 1 ˥	High Level 高平調	婚	夫	因	醫医	些	詩诗
Tone 2 ˧˥	High Rising 高升調	粉	苦	隱隐	倚	寫写	史
Tone 3 ˧	Mid Level 中平調	訓训	富	印	意	卸	嗜
Tone 4 ˩˧	Low Falling 低降調	焚	扶	人	兒儿	蛇	時时
Tone 5 ˨˩	Low rising 低升調	奮奋	婦妇	引	耳	社	市
Tone 6 ˨	Low level 低平調	份	負	孕	二	射	事

Table 1. Six base syllables having valid lexical items for each of the six long tones in Hong Kong Cantonese. Both the written forms and the Jyutping transcriptions are shown.

(2) Speech stimuli used in the experiment

(a) Training materials of large variability

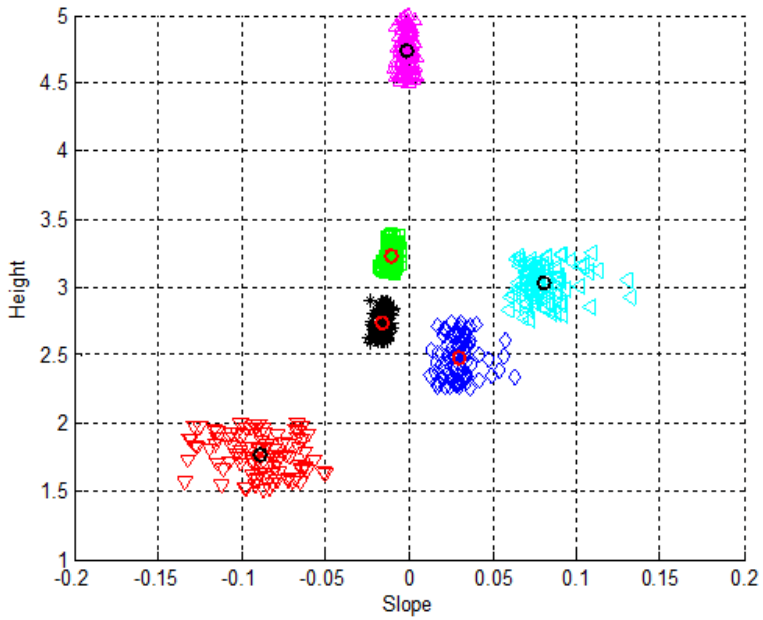


Figure 4. The tone chart for all the speech utterances used as training materials of large variability.

(b) Training materials of small variability

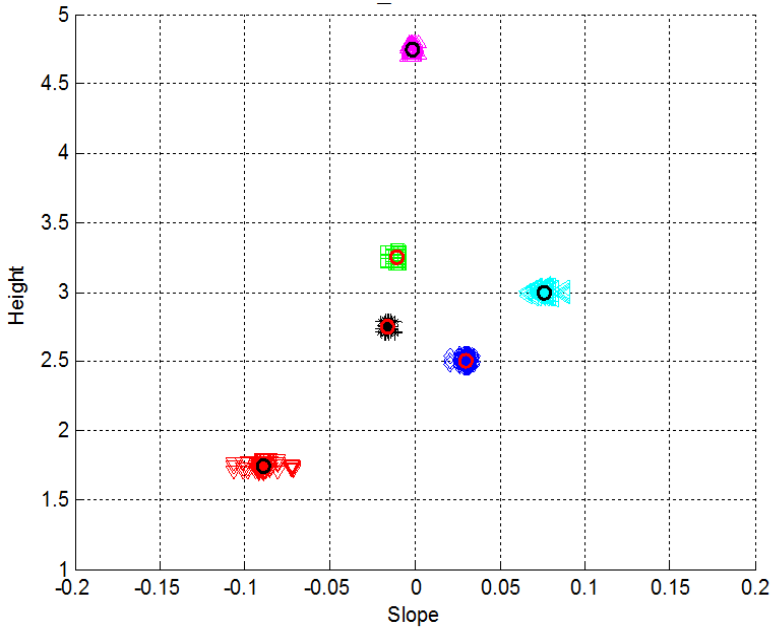


Figure 5. The tone chart for all the speech utterances used as training materials of small variability.

(c) The test stimuli

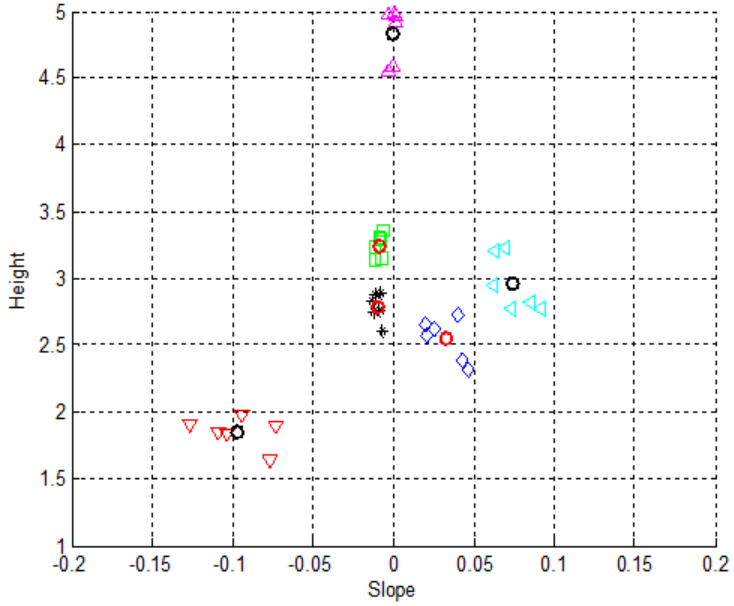
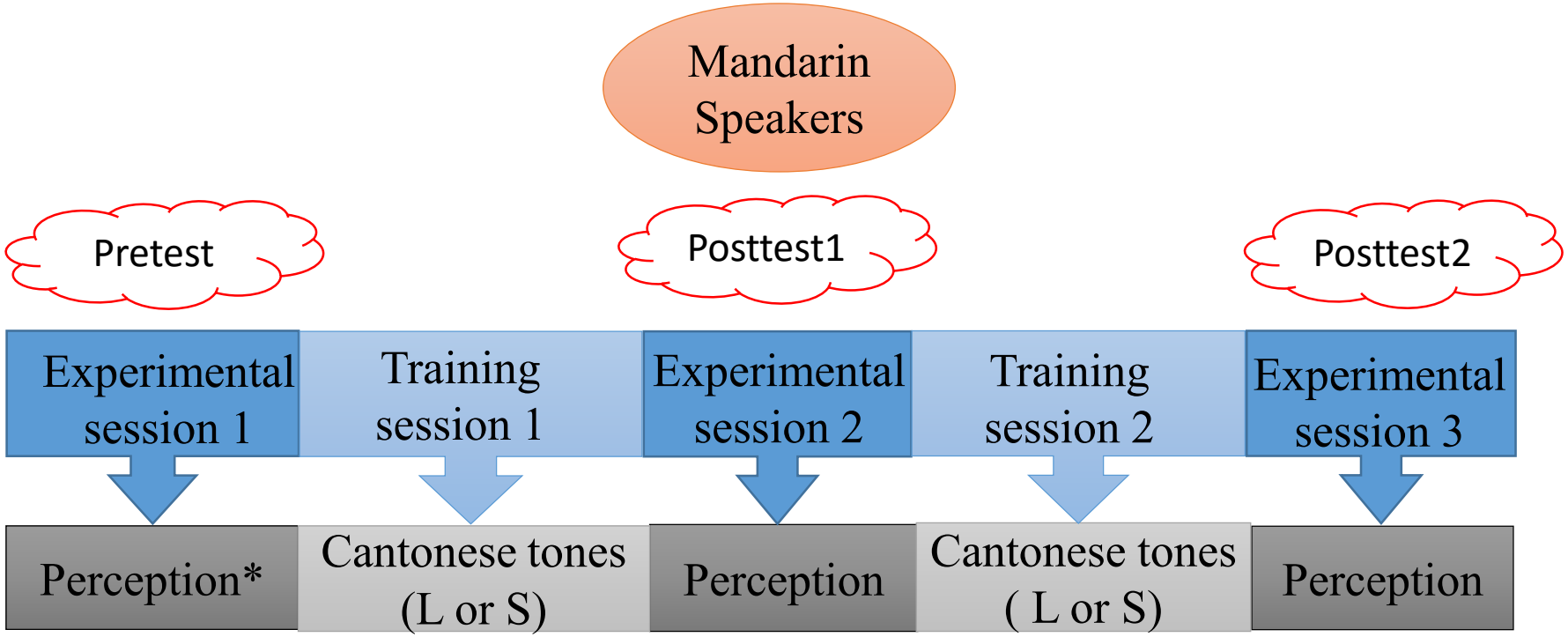


Figure 6. The tone chart for all the speech utterances used as test stimuli.

2.3 Procedures

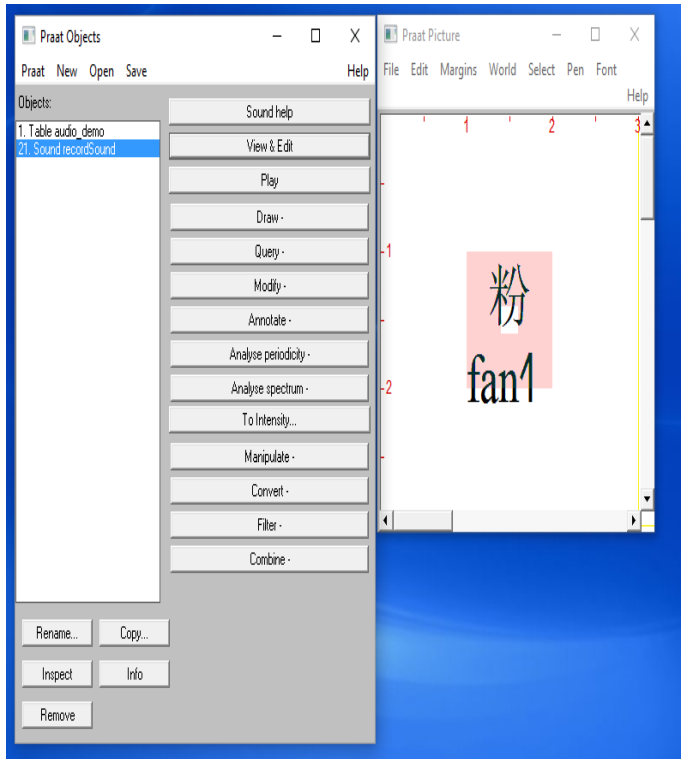


* Perception: Cantonese tone identification task

(1) Training sessions:

- Subject Group L → Training materials of LARGE variability
  Subject Group S → Training materials of SMALL variability

- Speech shadowing: Repeating the model speech after hearing it.



(2) Experimental sessions: Cantonese tone identification task

- To identify the tone category of each syllable

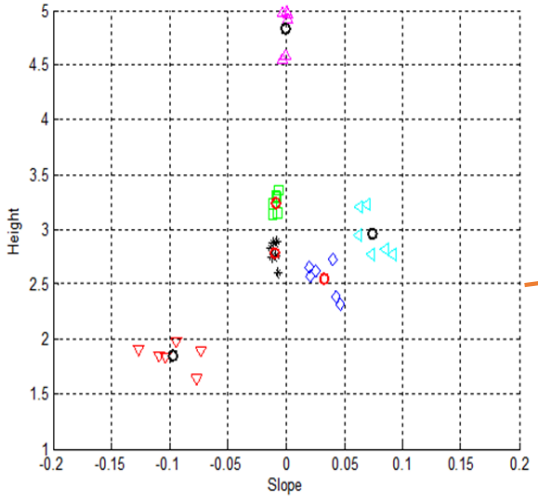


Figure 6. The tone chart for all the speech utterances used as test stimuli.

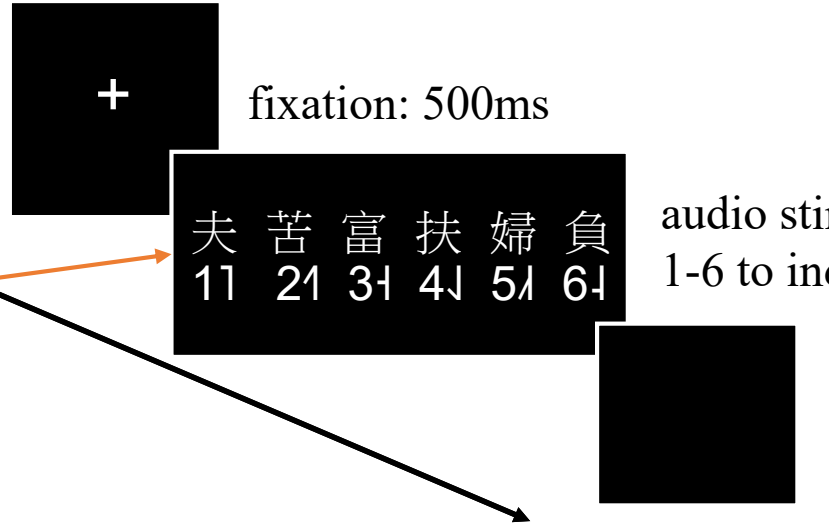


Figure 7. The illustration of the trail procedure in the Cantonese tone identification task.

3. Results

3.1 The accuracy analysis of tone identification task (a three-way ANOVA)

❑Between-subject factor:

- *Group* (*Ps* > 0.05)

❑Within-subject factors:

- *Experimental session*:

  S1 0.623 < S2 0.685 ≈ S3 0.686

- *Tone*:

  T21 (0.952) > T55 (0.899) > T33 (0.635) ≈ T23 (0.634) ≈ T25 (0.588) > T22 (0.279)

- *Base syllable*:

  ji (0.707) ≈ fu (0.702) > si (0.675) ≈ fan (0.664) > se (0.62) ≈ jan (0.619)

3.2 Confusion matrixes analysis

(a) Mandarin Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	1%	0%	2%	7%	90%
Target_33	1%	9%	1%	1%	64%	24%
Target_25	1%	0%	39%	59%	0%	1%
Target_23	1%	1%	64%	32%	1%	1%
Target_22	1%	28%	2%	1%	56%	12%
Target_21	95%	2%	1%	1%	1%	0%

(b) Cantonese Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	0%	0%	1%	1%	98%
Target_33	1%	14%	2%	1%	82%	1%
Target_25	1%	1%	22%	76%	1%	1%
Target_23	2%	1%	79%	17%	1%	0%
Target_22	1%	49%	1%	1%	47%	1%
Target_21	98%	1%	2%	0%	0%	0%

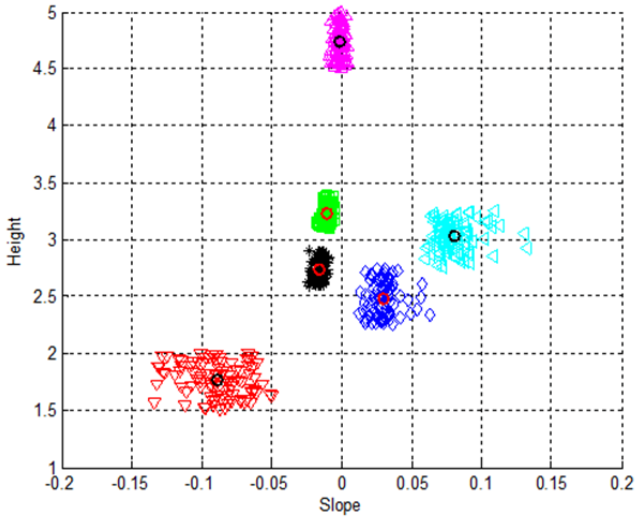
Table 2. Confusion matrix of tone identification for (a) Mandarin and (b) Cantonese participants.

- ❑ Tones of the similar slope are easily confused with each other.
  - T25/T23; T55/T33/T22
- ❑ Mandarin speakers show different confusion patterns from Cantonese speakers, especially for the two level tones that are not in Mandarin.
  - Can. (T33→ T22; T22→ T33)
    Man. (T33→ T55; T22→ T33/T55)

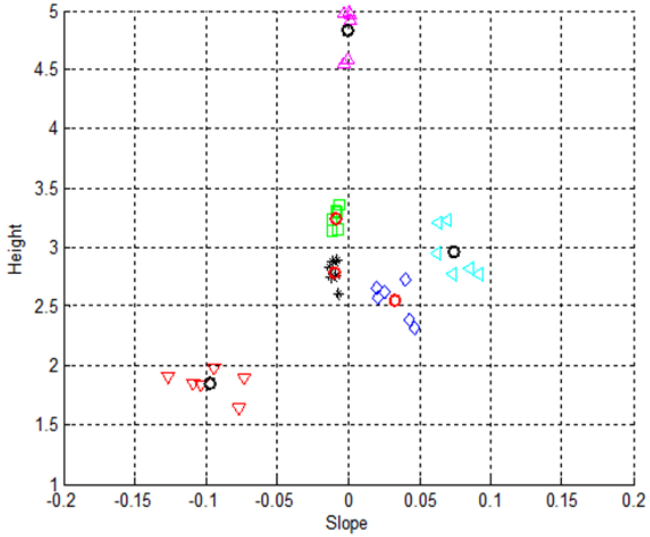
4. Discussion

4.1 the cognitive representation and perception of phonological categories

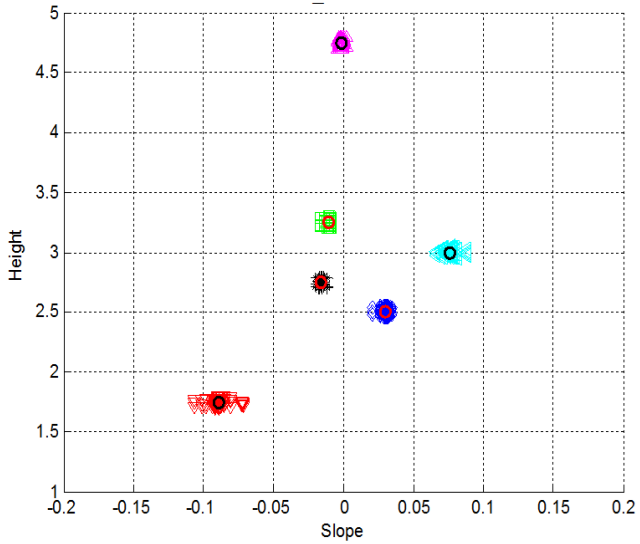
Large variability



The test stimuli



Small variability



- Group Large ≈ Group Small
- Exemplar-based model (×): not a simple match between the income signals and the exemplars stored mentally

4.1 the cognitive representation and perception of phonological categories

The abstract model (e.g., Gerstman, 1968; Syrdal & Gopal, 1986)

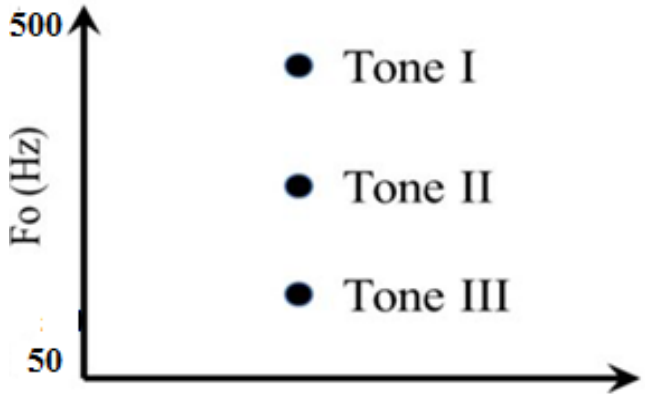


Figure 8. Summary of the abstract model, exemplified for a system of three tones.

The hybrid model (e.g., Goldinger, 2007)

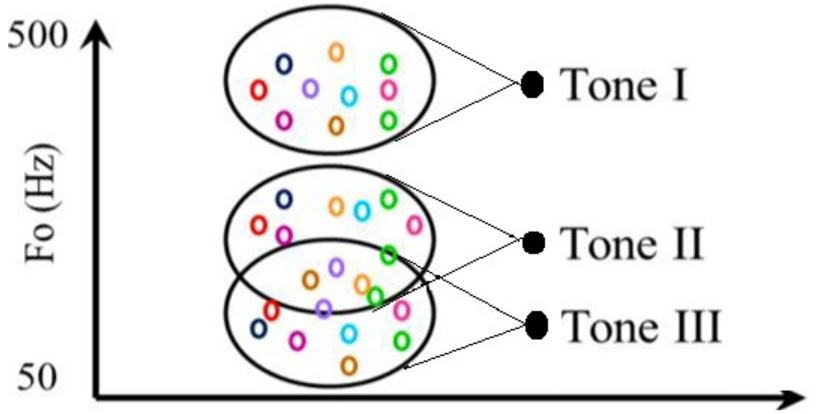


Figure 9. Summary of the hybrid model, exemplified for a system of three tones.

- A cognitive process in which the talker-specific variation is filtered out or gradually decays must occur.

4.2 The influence of L1 (i.e., Mandarin) background on learning Cantonese tones

(a) Cantonese Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	0%	0%	1%	1%	98%
Target_33	1%	14%	2%	1%	82%	1%
Target_25	1%	1%	22%	76%	1%	1%
Target_23	2%	1%	79%	17%	1%	0%
Target_22	1%	49%	1%	1%	47%	1%
Target_21	98%	1%	2%	0%	0%	0%

(b) Mandarin Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	1%	0%	2%	7%	90%
Target_33	1%	9%	1%	1%	64%	24%
Target_25	1%	0%	39%	59%	0%	1%
Target_23	1%	1%	64%	32%	1%	1%
Target_22	1%	28%	2%	1%	56%	12%
Target_21	95%	2%	1%	1%	1%	0%

(1) The nonnative speech signal is not identified directly but modulated by the native Mandarin phonological system first.

- Can. : T33 as T22 (acoustically closer)
  Man. : T33 as T55 (existing in the Mandarin)
- the Perceptual Assimilation Model (So &Best, 2010; 2014)

4.2 The influence of L1 (i.e., Mandarin) background on learning Cantonese tones

(a) Mandarin Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	1%	0%	2%	7%	90%
Target_33	1%	9%	1%	1%	64%	24%
Target_25	1%	0%	39%	59%	0%	1%
Target_23	1%	1%	64%	32%	1%	1%
Target_22	1%	28%	2%	1%	56%	12%
Target_21	95%	2%	1%	1%	1%	0%

(b) Cantonese Participants

	Real_21	Real_22	Real_23	Real_25	Real_33	Real_55
Target_55	0%	0%	0%	1%	1%	98%
Target_33	1%	14%	2%	1%	82%	1%
Target_25	1%	1%	22%	76%	1%	1%
Target_23	2%	1%	79%	17%	1%	0%
Target_22	1%	49%	1%	1%	47%	1%
Target_21	98%	1%	2%	0%	0%	0%

(2) Mandarin speakers might be more sensitive to the change of the pitch direction than the change of the pitch height.

- Tones with similar slope (i.e. three level tones and two rising tones) are more easily confused with each other.
- the language-specific weighting on pitch height and pitch slope (Gandour, 1983)

the Mandarin tonal system (T55, T35, T214 & T51: different pitch slopes)

4.2 The influence of L1 (i.e., Mandarin) background on learning Cantonese tones

(3) The acquisition of the tonal system in L2 is also affected by the segmental components.

- ji (0.707) ≈ fu (0.702) > si (0.675) ≈ fan (0.664) > se (0.62) ≈ jan (0.619)
- Mandarin &Cantonese: ji /ji/ & fu /fu/
  Cantonese: si /si/, fan /fɐn/, se /sɛ/ & jan /jɐn/
- Cognitive resource

5. Conclusion

(1) The original exemplars of a phonological category we encountered might undergo some changes and finally form the mental representations.

(2) The formation of a new tonal system is affected by the L1 phonological system from the suprasegmental to the segmental level.

6. Selected references

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149–175.

Gerstman, L. (1968). Classification of self-normalized vowels. *IEEE Transactions of Audio Electroacoustics*, AU-16: 78-80.

Goldinger, S.D. (1996). Words and Voices: Episodic Traces in Spoken Word Identification and Recognition Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166-1183.

Goldinger, S.D. (2007). A complementary-systems approach to abstract and episodic speech perception. In: Proceedings of the 2007 International Congress on Phonetic Sciences (pp. 49-54). Saarbrucken, Germany.

Johnson, K. (1997a). Speech perception without speaker normalization: An exemplar model. In K. Johnson, and J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 145-166). San Diego: Academic Press.

Johnson, K. (1997b). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics* 50: 101-113.

Peng, G. (2006). Temporal and tonal aspects of Chinese syllables: A corpus-based comparative study of Mandarin and Cantonese. *Journal of Chinese Linguistics* 34:134-154.

Peterson, G.E. & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175–184.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In: Bybee, J. and P. Hopper (eds) Frequency effects and the emergence of linguistic structure (pp. 137-157). John Benjamins, Amsterdam.

Syrdal, A.K., & Gopal, H.S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79: 1086-1100.

So, C. K. and 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.

So, C. K. and Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition* 36: 195–221.

Acknowledgments

The study is supported in part by a grant from Research Grant Council of Hong Kong (GRF: 14411314).

Thank you!