

The Discrimination of English Vowels by Cantonese ESL Learners in Hong Kong: A Test of the Perceptual Assimilation Model

Alice Y. W. Chan

City University of Hong Kong, Hong Kong, China
Email: enalice@cityu.edu.hk

Received January 6th, 2013; revised March 1st, 2013; accepted March 9th, 2013

Copyright © 2013 Alice Y. W. Chan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article discusses the results of a study which investigated Cantonese ESL learners' perception of English vowels and their perceived similarity between similar L1 and L2 vowels in an attempt to test the prediction of the Perceptual Assimilation Model (PAM). Forty university English majors participated in three L2 perception tasks, which aimed at discerning their perception of English vowels spoken in different contexts, and one L1 L2 speech perception task, which aimed at discerning their classification of L2 vowels into native vowel categories and their perceived similarity between similar L1 and L2 vowels. It was found that their classifications of English vowels into Cantonese vowels and their perception of the corresponding English vowels did not provide strong support for the prediction of the model. The effects and extent of native language phonological influence are yet to be determined.

Keywords: Second Language Acquisition; Speech Perception; Phonetics and Phonology

Introduction

A lot of research into second language phonology acquisition is centered around speech production, and mother tongue influence has often been argued as one major contributor to learner difficulties, in the sense that L2 sounds which are different from the L1 sounds are often difficult to produce. Mother tongue influence is, however, prevalent not just in the speech production arena, as research also shows that it has tremendous effects on the perception of L2 speech sounds, albeit in a different fashion. Flege (1995), for example, argues in his Speech Learning Model (SLM) that the more similar an L2 sound is to an L1 sound, the more problems an L2 learner will have in perceiving the L2 sound, because L2 learners are likely to judge L2 sounds as realizations of an L1 category. If L2 learners can detect the phonetic differences between an L2 sound and the nearest L1 sound, then they can perceive the L2 sound more easily. If not, problems will arise. Similarities, rather than differences, between the native and target languages are thus seen as the main contributor to learner difficulties. Another well-known model which attributes L2 learners' discrimination problems to the phonetic similarity between L1 and L2 sounds is the Perceptual Assimilation Model (PAM), to which the focus of the present article will turn.

Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM), developed by Best (1994), proposes that non-native contrasts are perceived in terms of their phonetic similarity to the phonological categories present in a listener's native language (Harnsberger, 2001). It

posits that "non-native speech perception is strongly affected by listeners' knowledge (whether implicit or explicit) of native phonological equivalence classes, and that listeners perceptually assimilate non-native phones to native phonemes whenever possible, based on detection of commonalities in the articulators, constriction locations and/or constriction degrees used" (Best, 1993; cited in Best, McRoberts, & Goodell, 2001: p. 777). The similarity between the native and target languages is seen as a vital factor determining L2 speech perception, as the degree of gestural similarity determines the matching between native phoneme categories and non-native phones. A listener will not be able to detect discrepancies between native and non-native phonemes if he or she perceives the nonnative phones to be very similar to a native phoneme category in their articulatory-gestural properties.

Best, McRoberts and Sithole (1988) (cited in Best, 1994) have listed four patterns of assimilation, which can be used to predict how well listeners will discriminate different foreign sounds from one another.

- 1) TC (Two Categories): The members of a non-native contrast may be gesturally similar to two different native phonemes, thereby assimilated to two categories;
- 2) SC (Single Category): The non-native phones may assimilate equally well, or poorly, to a single native category;
- 3) CG (Category Goodness): The non-native contrasts may both be assimilated to a single native category, with one more similar than the other to the native phoneme; and
- 4) NA (Non-assimilable): The non-native sounds may be too discrepant from the gestural properties of any native categories to be assimilated into any categories of the native phonology.

These should be perceived as non-speech sounds.

According to the PAM, only some non-native contrasts are difficult for mature listeners (phonologically sophisticated listeners) to discriminate, whereas others should be easy to discriminate even without prior training or exposure. The discrimination performance pattern for adults from highest performance to lowest should be: TC > (NA < = > CG) > SC (Best, 1994). Such a prediction assumes strong phonological influence from the L1, and the perceptual variations depend on the differences in the gestural similarities and discrepancies between the non-native contrasts and the native phonemes. For NA contrasts, discrimination performance depends on how similarly the two sounds are perceived to be non-speech sounds. It was claimed, in Best (1994), that the pattern of performance they obtained with adult listeners across several experiments with non-native speech contrasts had been consistent with this prediction. Other research studies carried out by Best and her collaborators also support this prediction (e.g. Best, McRoberts, & Goodell, 2001).

Current Research into the PAM

Since the introduction of the PAM, a number of research studies have been carried out to test their proposals and/or to investigate L2 or foreign language learners' speech perception abilities. Aoyama (2003), for example, investigated Korean and Japanese speakers' perception of English nasals to examine how learners' L1 influenced the perception of L2 segments. It was found that the speakers' performance was consistent with the prediction of the PAM: The final /n/-/ŋ/ contrast was particularly difficult, because neither sound was consistently classified with one L1 category and the same L1 categories were used for both. On the other hand, Kingston (2003) obtained data incompatible with the claims of the PAM in his study of the ability of American English learners to categorize German non-low vowels: He found that pairs of vowels contrasting minimally for the same feature in German often would not assimilate in the same way to English vowels, so some instances of the same contrast between German vowels were more easily discriminated than others. The ease with which a learner could tell one non-native phoneme from another, thus, did not vary directly with the extent to which these sounds assimilated to different native phonemes. In his investigation of the production and perception of Australian English vowels by Vietnamese and Japanese ESL speakers, Proctor (2004) also argued that although the PAM was useful at explaining some aspects of L2 phonology, there was a need for a more unified approach which could account for other issues such as temporal transfer (the transfer of skills in the perception of duration). Other research studies which claimed to have found supporting evidence for the assertions or basic premises of the PAM include Imsri (2003), who found that inexperienced learners perceived non-native sounds according to their L1 inventory; and Pilus (2002), whose data pointed to learners' better perception abilities than production abilities. Those which raised problems for the PAM or implicate factors other than perceptual assimilation include Harnsberger (2001), who argued that discriminability of non-native contrasts was a function of the similarity of non-native sounds to each other in a multidimensional, phonologized perceptual space; and Strange, Akahane-Yamada, Kubo, Trent, Nishi and Jenkins (1998) and Strange, Akahane-Yamada, Kubo, Trent and Nishi (2001), who argued that identification and dis-

crimination of L2 vowels varied significantly as a function of the contexts in which they were produced and presented.

Phonology Acquisition by Cantonese ESL Learners in Hong Kong

Many research studies have been carried out to investigate Cantonese ESL learners' second language phonology acquisition, most of which focus on learners' difficulties in the production of English speech sounds (e.g. Bolton and Kwok, 1990; A. Y. W. Chan, 2006a, 2006b, 2007; C. Y. H. Chan, 2005, 2007; Chan & Li, 2000; Hung, 2000, 2005; Lo, 2007; Stibbard, 2004). Both segmental problems (including problems in vowels, in consonants and in consonant clusters) and suprasegmental problems (such as word stress and rhythm) have been documented. With regard to the production of English vowels, substitution by a near sound in the native language has been reported as a most common strategy used to cope with problematic English sounds non-existent in the L1. For example, English /ɜ:/, a short vowel not found in Cantonese, is often replaced by a similar Cantonese vowel in production, namely /e/, as in words such as *leng3* /leŋ/ (The number at the end of each Cantonese word is a tone mark indicating one of the nine distinctive tones in Cantonese). English tense and lax vowel pairs, such as /i:/ and /ɪ/, /u:/ and /ʊ/, and /ɔ:/ and /ʌ/, have often become indistinguishable in length in the speech of Cantonese ESL learners. "Depending on individual learners, some may use a short vowel for a long one, others a long vowel for a short one; still others may produce a vowel sound which is somewhere in between the long and short vowels when pronouncing either one" (Chan & Li, 2000: pp. 80-81; see also Stibbard, 2004). Other widespread mispronunciation features include the unnecessary lip-rounding in the production of the central vowel /ʌ:/ (e.g. in words such as *bird*) and the substitution of pure vowels for diphthongs (e.g. /ɛ:/ for /aʊ/ in words such as *time*). These problems in producing English vowels are often explained in terms of the inventory gaps between the two languages, that L2 sounds non-existent in the native language are more difficult than those shared by both the native and target phoneme inventories, and that the substitution sounds often bear some articulatory and acoustic resemblance to the closest L1 sounds.

Research into the perception of English speech sounds by Cantonese ESL learners in Hong Kong has, to the author's knowledge, been very scarce. Chan (2001) is one notable exception. She studied Cantonese ESL learners' perception of English word-initial consonants and found a positive correlation between perception problems and production problems: Learners who consistently demonstrated perceptual confusion for the contrast pairs (/v, w/, /ʃ, f/, /tʃ, dʒ/, /z, s/ and /r, w/) also demonstrated confusion in production, and the target items /v, ʃ, tʃ, z, r/ were often misperceived as the same as their mispronounced versions /w, f, d, s, w/ respectively. Chan (2001) explained the results in terms of Bradlow et al. (1997)'s suggestion, that there might be a common mental representation determining both speech perception and speech production. Her data also supported Flége's (1991, 1992) model of L2 speech learning, that "L2 learners tend to perceive L2 sounds categorically within the sound classes of their L1" (Chan, 2001: p. 39). Another study which incorporated speech perception is Hung (2000). In this study of the phonology of Hong Kong English, Hung conducted a perception test of English vowels and found that his subjects could not distinguish pairs of vowels such as /i:/ and

/ʊ/, and /ɔ/ and /e/. The focus of the study, however, was on production, and the perception tests were just meant to provide further support for the production data obtained in the study and the acoustic analyses given rather than to investigate learners' perceptual abilities. No systematic research, to the author's knowledge, has been carried out to investigate the perception of English vowels by Cantonese ESL learners in Hong Kong, nor has there been any attempt to attribute learners' perception difficulties or abilities to established models such as the PAM. The present research, which is a sub-study of a large-scale project on the perception and production of L2 speech sounds by Cantonese ESL learners in Hong Kong, serves to bridge this research gap.

The Study

The present study examined ESL learners' perception of L2 vowels and their perceived relations between L1 and L2 vowels with an aim to investigate the extent to which the prediction of the PAM regarding different pairs of non-native contrasts are valid for explaining second language phonology acquisition by Cantonese ESL learners in Hong Kong.

Participants

A group of forty Hong Kong ESL learners (all native speakers of Cantonese) participated in the study. Twenty-nine of them were females and eleven males, with ages ranging from nineteen to forty-two at the time of the study. They all studied English as their majors at three local universities, including eight year 1 students, twenty-two year 2 students, and ten year 3 or postgraduate students. All of them started to learn English formally at the age of six or earlier when they entered primary schools. Twenty-six claimed to have received some form of phonetics training (such as taking a phonetics and phonology or pronunciation course), and the accent they learnt was Received Pronunciation (RP) English. Fourteen had not received any phonetics training before.

Perceptual Targets and Procedures

Three L2 perception tasks and one L1 L2 perception task were conducted to investigate the participants' perception of L2 vowels and their perceived similarity between "similar" L1 and L2 vowels. A total of eight English vowels, including three long and short vowel pairs, namely, /i:/, /ɪ/, /u:/, /ʊ/, /ɔ:/, /ɒ/, and the vowel pair /ɔ:/, /ə/, were included in all the L2 perception and L1 L2 perception tasks. Cantonese vowels which have "similar" acoustic and articulatory features with a target English vowel (e.g. Cantonese /i/ with English /i:/ and /ɪ/) were included for contrast in the L1 and L2 perception task. The English stimuli were spoken in RP English and the Chinese stimuli were spoken in Cantonese. The stimuli were presented to the participants individually at a comfortable volume over earphones in a quiet room during the implementation and a research assistant was responsible for administering the experiments.

L2 Categorical Discrimination Task (Task 1)

A categorical AXB discrimination test based on Best, McRoberts and Goodell (2001) was conducted to investigate the participants' discrimination of phones in isolation. In this task, series

of three isolated phones (i.e. AAB (e.g. u:, u:, ʊ), ABB (e.g. u:, ʊ, ʊ), BBA, (e.g. ʊ, ʊ, u:) or BAA (e.g. ʊ, u:, u:)) were presented. The participants were given a response sheet with a list of AXB sequences and asked to listen to the recorded stimuli and determine for each series whether the middle item (X) was the same as the first or the third item.

Word Discrimination Task (Task 2)

The purpose of the second task was to test the participants' ability to differentiate English minimal pairs. English words (e.g. *fool*) were spoken in isolation. A response sheet with the recorded word (e.g. *fool*) and a word differing in only one phoneme (e.g. *full*) was given to the participants. They had to listen to the recording and indicated the word they had heard from the corresponding pair on the response sheet (see **Appendix 1** for some sample pairs of words).

Picture Discrimination Task (Task 3)

The picture discrimination task tested the participants' ability to differentiate English minimal pairs spoken in carrier sentences (e.g. *Now I say _____*). A response sheet with a picture showing the recorded word (e.g. *pool*) and another picture showing a word in a minimal pair relationship with the recorded word (e.g. *pull*) was given to the participants, who had to indicate the picture which showed the word they had heard (see **Appendix 1** for some sample pairs of words).

Classification of English Vowels into Cantonese Vowels and Rating of Similarity (Task 4)

The task was divided into two parts. In the first part, a set of English words spoken in RP English and corresponding Cantonese words with "similar" vowels were presented to the participants. They had to classify the target English vowel (e.g. /ɔ:/) as a Cantonese vowel when hearing an English word and its corresponding Chinese list. For instance, when hearing an English stimulus [kʊk] *cook*, the participants had to classify the English vowel /ɔ:/ as one of the Cantonese vowels in a given list of Chinese words spoken in Cantonese (e.g. [kək] *kok8*, [kuk] *kuk7*, [kek] *kek9*, [kæk] *koek8*). The target English word (e.g. *cook*) was then presented to the participants for a second time, who had to rate the English vowel (e.g. /ɔ:/) for the degree of similarity to the Cantonese vowel just selected (e.g. /u/ in [kuk] *kuk7*) using a scale ranging from 1 (very different) to 5 (very similar). These two parts of the task required the participants to give both a classification response and a goodness-of-fit rating before proceeding to the next set of words. No previous training was provided for either the classification task or the rating task, but the participants were given a written list of all the words spoken (see **Appendix 2** for some sample sets of words).

Data Analysis

For Tasks 1 to 3, the proportion of correct judgments by the participants on each English sound and/or sound pair was computed to reveal the frequency with which a particular Eng-

¹A Proportion Z Test is a test of differences between two proportions from independent samples. Assuming that the samples are normally distributed, if Z (Z -value) > 1.96 , then there is a significant difference between the two proportions at the .05 significance level. Otherwise, the difference can be attributed to sampling errors.

lish sound or sound pair was correctly perceived in each task and in all the tasks. Proportion Z Tests¹ were conducted to determine the significance of the differences between the participants' performance on different sound pairs or on individual members of a pair.

For the first part of Task 4, the percentage of times that a particular English phone (e.g. /ɥ/) was classified as instances of a Cantonese sound category (e.g. /u:/, /e/, /i/, etc.) was computed. Classification overlap scores (Flege and Mackay, 2004) were also calculated for each pair of English contrasts. For example, if the participants had classified English /u:/ and /ɥ/ as Cantonese /u/ for p% and q% of instances respectively, then the classification overlap was q% if p > q, but p% if p < q.

For the second part of Task 4, the perceived similarity between a pair of L1 and L2 vowels (e.g. Cantonese /ɥ/ and English /u:/; Cantonese /i/ and English /i:/) was found by computing the mean goodness-of-fit rating that the participants assigned to the pair. For each degree of similarity ranging from 1 to 5, the product of the degree of similarity and the number of participants who chose that degree was first computed, then all the products were added up, and the sum was divided by the total number of participants.

Statistical analyses were conducted by SPSS 14.0. T-Tests were run to compare the mean goodness-of-fit ratings between different pairs of English and Cantonese contrasts (e.g. English

/u:/ and Cantonese /u/ with English /ɥ/ and Cantonese /u/).

Results

L2 Perception Tasks (Tasks 1-3)

Table 1 shows the participants' perception of different English vowels in different tasks. It can be seen that their perception was generally good. About 76% of all the target vowels were accurately perceived. Their perception of the vowel pair /ɔ:/, ①/ was the poorest. Overall accuracy rate was only 69%. (76% for /ɔ:/ and 62% for /①/). More instances of /①/ were misperceived as /ɔ:/ than vice versa. The /ɥ, e/ pair also presented a number of perceptual problems to the participants, with an overall accuracy rate of 77% and a similar number of both sounds accurately perceived (75% for /ɥ/ and 79% for /e/). The accuracy rates for /u:/ and for /ɥ/ were similar (76% and 79% respectively). /i:, ɥ/ was the best pair of vowels for perception. 81% of these vowels were accurately perceived, but the accuracy rate for /i:/ was only 73% whereas that for /ɥ/ was 90%. When individual members of tense and lax vowels were compared, it can be seen that lax vowels were on the whole more accurately perceived than corresponding tense ones. The only exception was /ɔ:, ①/.

Table 1.
Perception of different vowel pairs by the participants.

Vowels	Task 1	Task 2	Task 3	All Tasks	Z-statistics between first member and second member
Percentages of sounds correctly perceived					
i:, ɥ	N = 160	N = 160	N = 200	N = 520	Z = 4.91*
	92%	89%	65%	81%	
i:	N = 80	N = 80	N = 120	N = 280	
	98%	89%	46%	73%	
ɥ	N = 80	N = 80	N = 80	N = 240	Z = .78
	86%	90%	93%	90%	
u:, ɥ	N = 160	N = 160	N = 160	N = 480	
	98%	75%	59%	77%	
u:	N = 80	N = 80	N = 40	N = 200	Z = 3.43*
	98%	68%	48%	76%	
ɥ	N = 80	N = 80	N = 120	N = 280	
	99%	83%	63%	79%	
ɔ:, ①	N = 160	N = 160	N = 200	N = 520	Z = 1.12
	99%	69%	44%	69%	
ɔ:	N = 80	N = 80	N = 80	N = 240	
	100%	89%	40%	76%	
①	N = 80	N = 80	N = 120	N = 280	Z = 3.43*
	99%	50%	46%	62%	
ɥ, e	N = 160	N = 200	N = 200	N = 560	
	100%	75%	61%	77%	
ɥ	N = 80	N = 80	N = 80	N = 240	Z = 1.12
	100%	68%	56%	75%	
e	N = 80	N = 120	N = 120	N = 320	
	100%	80%	63%	79%	
Average	N = 640	N = 680	N = 760	N = 2080	
	97%	77%	57%	76%	

*Difference is significant at the .05 level.

Proportion Z tests showed that the difference between /i:/ and /ɨ/ and that between /ɔ:/ and /ɔ̃/ were significant at the .05 significance level, whereas the difference between /u:/ and /ʊ/ and that between /ɔ̃/ and /e/ were non-significant (see Z-statistics in **Table 1**). Proportion Z tests also showed that the differences in overall accuracy rates between the /ɔ:/, /ɔ̃/ pair and all other pairs were significant, whereas the differences in overall accuracy rates between other pairs of vowels were not statistically significant at all (not shown in **Table 1** to avoid confusion).

Classification of English Vowels as Cantonese Vowels (Task 4a)

Table 2 shows the participants' classification of English vowels as Cantonese vowels. English /i:/ and /ɨ/ were predominantly classified as Cantonese /i/ (91% and 89% respectively), English /u:/ and /ʊ/ as Cantonese /u/ (96% and 93% respectively), and English /ɔ:/ and /ɔ̃/ as Cantonese /ɔ/ (88% and 95% respectively). This shows that all English tense and lax vowel pairs were predominantly classified as the "nearest" Cantonese lax vowels, which presumably have the closest articulatory features, i.e. Cantonese /i/, like English /i:/ and /ɨ/, is a high front vowel; Cantonese /u/, like English /u:/ and /ʊ/, is a high back vowel; and Cantonese /ɔ/, like English /ɔ:/ is at the mid back region³. Despite such predominant classifications, all the target English tense and lax vowels were also classified as other Cantonese vowels with rather different articulatory features. For example, 5% and 6% of English /i:/ and /ɨ/ respectively were classified as Cantonese /a/, whereas 8% of English /ɔ:/ were classified as Cantonese /u/. As for the /ɔ̃/, e/ pair, both

were predominantly classified as Cantonese /e/ (90% for /ɔ̃/ and 50% for /e/), though the latter showed more diverse classifications, with 13%, 19% and 18% being classified as Cantonese /i/, /a:/ and /a/ respectively.

Classification overlap was highest for /u:/, /ʊ/, with a score as high as 93%, and lowest for the low and mid vowel pair (/ɔ̃/, e/) (overlap = 50%). Overlap scores for other tense and lax vowel pairs were also high, with 89% for /i:/, /ɨ/ and 88% for /ɔ:/, /ɔ̃/.

Perceived Degrees of Similarity between English and Cantonese Vowels (Task 4b)

Table 3 shows the participants' perceived degrees of similarity between the target English vowels and the Cantonese vowels which they had selected as most similar, and **Table 4** shows the T-tests results. It can be seen that the mean goodness-of-fit ratings were mostly in the range between 3 and 4. For the English low and mid vowel pair, English /ɔ̃/ was regarded more similar to Cantonese /e/ than English /e/. The mean goodness-of-fit rating assigned for the former was 3.95 and that for the latter was 3.48. This difference was significant at the .05 significance level. The mean goodness-of-fit ratings assigned for English /i:/ and English /ɨ/ to Cantonese /i/ were 3.70 and 3.59 and those for English /u:/ and English /ʊ/ to Cantonese /u/ were 3.70 and 3.81 respectively. Neither of these differences between the corresponding English tense and lax vowels was statistically significant. The mean goodness-of-fit ratings for English /ɔ:/ and English /ɔ̃/ to Cantonese /ɔ/ were 3.59 and 4.01 respectively, and the difference between these two was statistically significant at the .05 significance level (see **Tables 3** and **4**).

Table 2.
Participants' classification of English vowels as Cantonese vowels.

Eng. vowels	Can. vowels	Percentages of English vowels classified as Cantonese vowels						
		i	a:	a	æ	u	ɔ	e
i:		91%	4%	5%	0%	0%		
ɨ		89%	3%	6%	3%	0%		
u:		0%	0%	0%	4%	96%	0%	
ʊ			0%	3%	1%	93%	4%	0%
ɔ:		0%	4%	1%		8%	88%	
ɔ̃			1%	1%	1%	1%	95%	0%
ɔ̃		5%	0%	4%		1%	0%	90%
e		13%	19%	18%		1%		50%

Table 3.
Participants' perception of degrees of similarity between English and Cantonese vowels.

Eng. vowels	Can. vowels	Mean goodness-of-fit ratings						
		i	a:	a	æ	u	ɔ	e
i:		3.70	3.00	3.00				
ɨ		3.59	4.00	3.20	2.50			
u:					2.67	3.70		
ʊ				1.50	4.00	3.81	3.00	
ɔ:			3.00	2.00		2.50	3.59	
ɔ̃			2.00	4.00	1.00	2.00	4.01	
ɔ̃		2.75		2.67		3.00		3.95

³In **Table 1**, the data are presented as results on a pair and results on individual items in the pair. If the accuracy rate of an individual item (e.g. /ɔ:/) is lower than 100%, then the misperceived tokens were perceived as the other item (e.g. /ɔ̃/) in the corresponding pair (e.g. /ɔ:/, /ɔ̃/).

e	2.70	2.40	2.36	3.00	3.48
---	------	------	------	------	------

Table 4.

Comparison of mean goodness-of-fit ratings for similar English and Cantonese vowels.

English and Cantonese Vowels	N	Mean	Mean Difference	Sig.
/i:/ and /i/	39	3.7051		
/ɨ/ and /i/	39	3.5897		
			.11538	.457
/u:/ and /u/	39	3.7051		
/ɯ/ and /u/	40	3.8125		
			-.10737	.578
/ɔ:/ and /ɔ/	39	3.5897		
/ɔ̃/ and /ɔ/	40	4.0125		
			-.42276	.014*
/ɛ/ and /e/	39	3.9487		
/e/ and /e/	40	3.4750		
			.47372	.007*

*Difference is significant at .05 level.

Interestingly, all the English lax vowels were classified by a minority of the participants as very similar (goodness-of-fit ratings = 4) to a “non-equivalent” Cantonese vowel: English /ɨ/ when compared with Cantonese /œ/, English /ɨ/ with Cantonese /a:/, and English /ɔ̃/ with Cantonese /a/ all received a goodness-of-fit rating of about 4.

Discussion

Prediction of the PAM

From the results of the study, it can be seen that some L2 vowel pairs were assimilated by the participants to a single native category with one more similar than the other to the native phoneme (CG: Category Goodness), and some should be seen as equally similar to a single native category (SC: Single Category). The /ɛ/ and /e/ pair was an example of the former. Both of these two sounds were classified as most similar to Cantonese /e/ with a low classification overlap of 50%, showing that this pair of non-native contrasts may have assimilated to Cantonese /e/ but with English /ɛ/ more similar than English /e/ to the native phoneme (CG). The statistically significant goodness-of-fit rating difference between the two (when compared to Cantonese /e/) also confirms that they should be regarded as CG. English /i:/, /ɨ/ and /u:/, /ɯ/ were good exemplars of the SC pattern, as they assimilated equally well to a single native category with classification overlaps as high as 89% or above. The statistically nonsignificant goodness-of-fit rating differences between the tense and lax vowels (when compared with the corresponding Cantonese vowel) also confirm this grouping. The English /ɔ:/, /ɔ̃/ pair, on the other hand, invites some indeterminacy in patterning. The classification overlap between the two was high (88%), suggesting that they should have been assimilated to a single category, but there was a statistically significant goodness-of-fit rating difference between the former and the latter when compared with the same Cantonese vowel /ɔ/, showing that they should be regarded as CG instead.

None of the English vowel pairs could be regarded as similar to two different native phonemes (TC: Two Categories): Although all the vowels were perceived by some participants as similar to some other L2 phonemes rather than the “nearest” one and the goodness-of-fit ratings were very high, the percentages of such classifications were too low to be of significance for comparison. As such, there was no TC pattern in the

study. NA (Non-assimilable) was not applicable in the study either.

As mentioned before, the PAM predicts that the discrimination performance pattern for adults from highest performance to lowest performance is TC > CG > SC. The participants’ classifications of English vowels into Cantonese vowels and their perception of the corresponding English vowels did not provide supporting evidence for this prediction. Their perception of the CG pair /ɛ/ and /e/ was largely the same as that of the SC pairs /u:/, /ɨ/ and /i:/, /ɯ/. Their perception of another CG pair /ɔ:/, /ɔ̃/ was actually the worst, statistically significantly poorer than their perception of the two SC pairs. With a performance pattern largely different from the pattern of the prediction, it is doubtful whether the prediction is substantiated and valid for explaining second language vowel acquisition by Cantonese ESL learners. The ease with which a Cantonese ESL learner can distinguish a pair of non-native contrast from another pair is, thus, not necessarily a function of the extent to which the L2 contrasts assimilate to the L1 phonemes.

Native Phonological Influence

Although the results of the present study do not give supporting evidence to the PAM’s predicted discrimination performance pattern, the model’s claims regarding native phonological influence and learners’ perception of non-native phones in terms of their L1 phonological categories are not to be falsified: Cantonese ESL learners do regard different English vowels as similar, albeit to different extents, to one or some of their native vowels, and the L1 phoneme prevalently perceived as similar to a certain L2 vowel is the one which shares the closest articulatory properties with the L2 sound, showing that learners do assimilate non-native phones to native phonemes based on detection of commonalities that exist between them in the articulations. The high goodness-of-fit ratings assigned to the L1 and L2 vowels also confirm learners’ perceived similarities between the L1 and the L2.

A notable pattern also seems to emerge from the results: The perception of an individual L2 sound bears an intimate relation with the perceived distance between the L2 sound itself and the closest L1 phoneme, asserting the basic premise of the PAM, that if a learner perceives a non-native phone as very similar to a native phoneme, he or she will not be able to detect discrepancies between the two. English /ɔ̃/, for example, which was

³Unlike English /ɔ:/ and Cantonese /ɔ/, English /ɔ̃/ is regarded as a low vowel rather than a mid vowel.

considered significantly more similar than English /ɔ:/ to Cantonese /ɔ/, was perceived significantly less accurately by the participants than English /ɔ:/ (accuracy rate for the former = 62% and that for the latter = 76%). The significantly smaller perceived similarity between English /e/ to Cantonese /e/ than English /ɔ/ to Cantonese /e/ also resulted in better perception of the former than the latter, although the difference was not statistically significant. Because vowels contrasting minimally in English do not assimilate in the same way to the same Cantonese vowel, a member (e.g. English /ɔ:/) of a vowel pair (e.g. English /ɔ:/, /ə/) may be more easily or difficultly perceived than the other (e.g. English /ə/). Rather than making reference to pairs of non-native contrasts and predicting learners' relative difficulty in perceiving one pair (e.g. CG) and another (e.g. TC) as the PAM does, it seems more appropriate to count on the perceived distance between an individual L2 vowel and a native phoneme. The more similar an L2 vowel is to an L1 vowel, the more difficultly the L2 vowel is perceived (Flege, 1995; also see Chan, 2012). With the limited pool of data obtained from the present study, no reliable conclusions can be drawn regarding perceived distance and perception. More research is needed to ascertain this relation as well as the extent of native phonological influence.

Conclusion

In this article, I have reported on the results of a research study which investigated the perception of English pure vowels by Cantonese ESL learners in Hong Kong in an attempt to test the PAM's prediction on discrimination performance. The results of the study do not provide strong support for the prediction, suggesting that a pair of non-native contrasts with one member classified as closer to a native phoneme than the other member may not necessarily be more accurately perceived than another pair which is classified as equally similar to a native phoneme category. Native language phonological influence is, however, not nullified in the area of speech perception. Rather than predicting perception performance with reference to pairs of non-native contrasts and learners' assimilation of these non-native contrasts to L1 phonemes, it seems more appropriate to predict L2 perception based on the perceived distance between a certain non-native sound and the closest native phoneme(s). The results of the study have theoretical significance, providing a platform for future research into the Perceptual Assimilation Model, the relationship between perceived similarity and L2 perception, as well as the extent of native phonological influence. As only a homogeneous group of tertiary-level participants participated in the study, the results cannot be generalized to all Cantonese ESL learners in Hong Kong, especially elementary learners. Further research is needed to include learners from different proficiency levels and, preferably, from speakers of other languages. It is also illuminating to include other non-native sound categories, such as consonants, as well as vowels in different phonological environments, if a full picture of native phonological influence is to be attained.

Acknowledgements

The work described in this article was fully supported by a competitive earmarked research grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project Number: CityU 1455/05H]. The support of the Council

is acknowledged. I would also like to thank the participants of the study for their contribution and my research assistant for her administrative assistance.

REFERENCES

- Aoyama, K. (2003). Perception of syllable-initial and syllable-final nasals in English by Korean and Japanese speakers. *Second Language Research*, 19.3, 251-265. doi:10.1191/0267658303sr222oa
- Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year* (pp. 289-304). Dordrecht: Kluwer Academic.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge: MIT Press.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775-794. doi:10.1121/1.1332378
- Best, C. T., McRoberts, G. W., & Sithole, N. N. (1988). The phonological basis of perceptual loss for nonnative contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345-360. doi:10.1037/0096-1523.14.3.345
- Bolton, K., & Kwok, H. (1990). The dynamics of the Hong Kong accent: Social identity and sociolinguistic description. *Journal of Asian Pacific Communication*, 1.1, 147-172.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299-2310. doi:10.1121/1.418276
- Chan, A. Y. W. (2006a). Cantonese ESL learners' pronunciation of English final consonants. *Language, Culture and Curriculum*, 19.3, 296-313. doi:10.1080/07908310608668769
- Chan, A. Y. W. (2006b). Strategies used by Cantonese speakers in pronouncing English initial consonant clusters: Insights into the interlanguage phonology of Cantonese ESL learners in Hong Kong. *International Review of Applied Linguistics in Language Teaching*, 44, 331-355. doi:10.1515/IRAL.2006.015
- Chan, A. Y. W. (2007). The acquisition of English word-final consonants by Cantonese ESL learners in Hong Kong. *Canadian Journal of Linguistics*, 52.3, 231-253. doi:10.1353/cjl.2008.0023
- Chan, A. Y. W. (2012). Cantonese English as a second language learners' perceived relations between "similar" L1 and L2 speech sounds: A test of the speech learning model. *The Modern Language Journal*, 96.1, 1-19. doi:10.1111/j.1540-4781.2012.01291.x
- Chan, A. Y. W., & Li, D. C. S. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, 13.1, 67-85. doi:10.1080/07908310008666590
- Chan, C. P. H. (2001). The perception (and production) of English word-initial consonants by native speakers of Cantonese. *Hong Kong Journal of Applied Linguistics*, 6.1, 26-44.
- Chan, C. Y. H. (2005). L1 and L2 phonological variation: The merging of the syllable-initial /n-/ with /l-/ in Cantonese and English by Hong Kong students. Paper presented at IACL 13, Leiden: Leiden University.
- Chan, C. Y. H. (2007). Factors affecting L2 pronunciation: The merging of the syllable-initial /n-/ with /l-/ by Cantonese speakers learning English. *The 32th Annual Congress of Applied Linguistics Association of Australia*. Wollongong: University of Wollongong.
- Flege, J. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Huebner, & C. Ferguson

- (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 249-289). Philadelphia: John Benjamins Publishing Company.
- Flege, J. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 565-604). Timonium: York Press.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore: York Press.
- Flege, J. E., & Mackay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1-34. doi:10.1017/S0272263104261010
- Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America*, 110.1, 489-503. doi:10.1121/1.1371758
- Hung, T. T. N. (2000). Towards a phonology of Hong Kong English. *World Englishes*, 19.3, 337-356. doi:10.1111/1467-971X.00183
- Hung, T. T. N. (2005). Word stress in Hong Kong English: A preliminary study. *Applied Language Studies*, 9, 29-40.
- Im Sri, P. (2003). The perception of English stop consonants by Thai children and adults. Doctoral Thesis, Newark, DE: University of Delaware.
- Kingston, J. (2003). Learning foreign vowels. *Language and Speech*, 46.2-3, 295-349. doi:10.1177/00238309030460020201
- Lo, S. K. (2007). The markedness differential hypothesis and the acquisition of English final consonants by Cantonese ESL learners in Hong Kong. M. Phil Thesis, Hong Kong: City University of Hong Kong.
- Pilus, Z. (2002). Second language speech: Production and perception of voicing contrasts in word-final obstruents by Malay speakers of English. Doctoral Thesis, Madison, WI: University of Wisconsin-Madison.
- Proctor, M. (2004). Production and perception of AusE vowels by Vietnamese and Japanese ESL learners. *2004 Australian Linguistic Society Annual Conference*. Sydney: University of Sydney.
- Stibbard, R. (2004). The spoken English of Hong Kong: A study of co-occurring segmental errors. *Language, Culture and Curriculum*, 17.2, 127-142. doi:10.1080/07908310408666688
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, 109.4, 1691-1704. doi:10.1121/1.1353594
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listener. *Journal of Phonetics*, 26, 311-344. doi:10.1006/jpho.1998.0078

Appendix 1

List of Word Pairs Used in Task 2 and Task 3

Task 2.

Minimal pair discrimination.

1) eat	it	2) fool	full	3) look	Luke
4) wok	walk	5) beg	bag	6) bin	bean
7) beach	bitch	8) pick	peak	9) suit	soot
10) hood	who'd	11) pod	pawed	12) don	dawn
13) cot	caught	14) bed	bad	15) sat	set
16) man	men				

Task 3.

Picture discrimination.

1) bean	bin	2) hit	heat	3) tin	teen
4) sit	seat	5) ship	sheep	6) look	Luke
7) full	fool	8) pool	pull	9) hood	who'd
10) caller	collar	11) wok	walk	12) chalk	choc
13) stock	stalk	14) not	nought	15) send	sand
16) bend	band	17) men	man	18) said	sad
19) pen	pan				

Appendix 2

Response Sheet for Task 4

Task 4.

Sample words only.

Instruction:

Below is a list of English words. You will hear each English word twice and each list of Chinese words once.

Task a) After hearing the English word and the list of Chinese words for the first time, classify the English vowel in the word as a Cantonese vowel. A Cantonese word for each given vowel has been supplied as hints.

Task b) After hearing the English word for the second time, rate the English vowel in the word for the degree of similarity to the Cantonese vowel you have just chosen, using the given scale ranging from 1 (very different) to 5 (very similar).

Set 1		
1	teen	a:(taan1), a(tan1), i(tin1), œ (teon5)
2	sit	a:(saat8), a(sat7), œ (seot7), i(sit8)
3	lip	u(luk9), a(lap9), a:(laap9), i(lip9)
4	beak	i(bik7), u(buk7), a(bak7), a:(baak7)
Set 4		
1	ket	e(kek9), u(kut8), i(kit8), a(kat7)
2	pack	a:(paak8), u(puk7), e(pek8), i(pik7)
3	men	a(man6), a:(maan6), i(min6), u(mun4)
4	bang	e(beng3), a(bang1), ɔ (bong1), i(bing1)