Title	The perception of English vowels by native Cantonese English as a foreign language primary school students
Author(s)	Pang, Hiu-wa; 彭曉華
Citation	Pang, H. [彭曉華]. (2014). The perception of English vowels by native Cantonese English as a foreign language primary school students. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2014
URL	http://hdl.handle.net/10722/238939
Rights	This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; The author retains all proprietary rights, (such as patent rights) and the right to use in future works.

The perception of English vowels by native Cantonese English as a foreign language primary school students

Pang Hiu Wa

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2014.

Abstract

This study investigated the perception of English vowels in Cantonese speakers who are beginning learners of English as a foreign language (EFL) in reference to the Perceptual Assimilation Model (PAM). Thirty-one primary school students participated in a perception study that required them to discriminate and identify English minimal vowel pairs. It was founded that Cantonese EFL learner's vowel perception can be predicted by the similarity of spatial proximity of constriction locations between English vowels and Cantonese vowels. The study also provides support for the PAM. Further research is needed to include EFL learners from different age groups and at different English proficiency levels.

Introduction

Numerous researches have claimed that non-native language perception is highly influenced by the listeners' first language (L1) (Bohn & Flege, 1990; Flege, 1995; Ingram & Park, 1997; Strange, Akahane-Yamada, Kubo, Trent, Nishi, & Jenkins, 1998; Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Best, McRoberts, & Goodell, 2001). They further explain that the ease or difficulty of learning a foreign language is mainly attributed to the impact of the L1 phonological knowledge. It is well known that acquisition of a foreign language can be challenging as the non-native phonetic segments are phonetically different or do not exist in their L1 (Best, 1994; Ingram & Park, 1997; Strange et al., 1998). Previous studies have shown that English as a foreign language (EFL) or second language learners usually have demonstrated difficulties in perceiving particular English vowels due to differences between English and their L1 phonological system (Flege, 1987, 1995; Bohn & Flege, 1990; Fox, Flege, & Munro, 1995; Flege, Bohn, & Jang, 1997; Ingram & Park, 1997; Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Chan, 2012;).

Various models of cross-language speech perception propose that the perceived relation between phonetic segments in a foreign language and L1 plays a significant role in how those non-native phonetic segments will be discriminated (Guion et al., 2000; Aoyama et al., 2004; Best & Tyler, 2007). The two influential approaches in the study of cross-language vowel perception are the Perceptual Assimilation Model (PAM) developed by Best (1995) and the Speech Learning Model (SLM) developed by Flege (1995).

The PAM is a direct realistic model which draws from articulatory phonological theory that listeners perceive information in speech regarding the articulatory gestures that produced the speech signal (Best, 1995; Best, et al., 2001). The major principle of the PAM of cross-language speech perception is that non-native phonetic segments are perceived according to their similarities of spatial proximity to the native segmental constellations in native

phonological space and assimilated to native categories (Best, 1995). Therefore, any perceived distance between the non-native speech segments and L1 segments leads to differences in discriminability. According to the PAM, there are three classifications of perceptual assimilation of non-native segment contrasts: 1) exemplar of native segment; 2) uncategorizable speech sound that falls within native phonological space; and 3) nonassimilable nonspeech sound. Perceptual difficulties in non-native vowel contrasts differentiation are predictable on the patterns of perceptual assimilation of non-native vowels to L1 categories. Different possible pairwise assimilation patterns for non-native vowel contrasts and their predicted discrimination levels are as follows: 1) Two-Category Assimilation (TC), the contrasting non-native vowels are assimilated to two different L1 categories, thus discrimination should be excellent as they are phonemically distinct; 2) Category-Goodness Difference (CG), the contrasting non-native vowels are assimilated to the same L1 category but differ in discrepancy from the L1 segment, where discrimination can range from fair to good depending on the degree of difference in category goodness for each vowel; 3) Single-Category Assimilation (SC), the contrasting non-native vowels are assimilated to the same L1 category and they are both equally discrepant from the L1 segment, thus discrimination should be poor; 4) Both Uncategorizable (UU), the contrasting non-native vowels both fall into the phonetic space that is outside the L1 categories, therefore discrimination can range from poor to excellent depending on the their proximity to each other and to the L1 categories; 5) Uncategorized versus Categorized (UC), one vowel is assimilated to a L1 category and the other one falls in phonetic space that is outside the L1 categories, thus discrimination should be excellent; 6) Nonassimilable (NA), both contrasting non-native vowels are detected as nonspeech sounds that fall outside of the speech domain, therefore discrimination can range from good to excellent depending on their discriminability

as nonspeech sounds (Best, 1995; Polka, 1995; Strange et al., 1998; Best et al., 2001; Escudero & Boersma, 2002).

Similar to the PAM, the SLM also addresses the significance of perceptual relationship between L1 and non-native phonetic categories. It posits that non-native segments are equivalence-classified relative to L1 segments according to the basis of phonetic similarity (Best et al, 2001). The more dissimilar the non-native segments are from the L1 closest segment, the more likely that new phonological categories can be established and more accurate perception and production can be achieved (Bohn & Flege, 1990; Flege, 1995; Guion et al., 2000; Chan, 2012). Flege (1995) further described that the SLM is mainly focused on the ultimate attainment of second language pronunciation of experienced second language learners, not beginners. There are only limited amounts of researches examined the perception of English vowels by native Cantonese EFL learners. Recently, a study by Chan (2012) reported that perception of English speech sounds is associated with perceived similarity between Cantonese and English. Participants in the Chan's (2012) study were advanced English as foreign language learners and therefore their performance could be related to the SLM. The results, however, cannot be generalized to most Hong Kong Cantonese speaking children who are beginning learners of EFL.

In this research study, the main focus was on the perception of English vowel contrasts by Cantonese EFL primary school students. Unlike the PAM, one of the shortfalls of the SLM is its focus on age-related learning of individual phonemes of a second language but does not provide clear predictions about the perception of non-native contrasts (Best et al., 2001; Escudero & Boersma, 2002). When the English proficiency of the participants in the current study was concerned, the main focus of the PAM seemed to be more relevant. The PAM was developed to explain the non-native speech perception by beginners such as those in the early stages of foreign language acquisition, whereas the SLM concerned mainly

experienced second language learners (Flege, 1995; Guion et al., 2000; Edcudero & Boersma, 2004; Pilus, 2005; Best & Tyler, 2007). In previous research studies, the PAM has been shown to be applicable to the non-native perception of English, French, German, Japanese, Malay, Norwegian and Zulu by foreign listeners or inexperienced foreign language learners (Best, 1995; Polka, 1995; Strange et al., 1998; Best et al., 2001; Pilus, 2005). In the current study, the perception of English vowel contrasts was examined by using a discrimination task and a word identification task. The application of the PAM in predicting the performance of EFL learners in Hong Kong was also evaluated.

Hypothesis

Based on the phonemic inventory of English and Cantonese and the classification of non-native segments in the PAM, English vowels are assimilated as categorizable speech sounds by native Cantonese EFL listeners (See Appendix A for the overview of English and Cantonese vowels). Certain patterns of assimilation of the English vowels to a particular or a cluster of Cantonese vowel/s are expected, possible pairwise assimilation patterns include TC, CG and SC (See Table 1 for the assimilation patterns and its prediction of discriminability). UU, UC and NA are inapplicable as they involve at least one non-native segment in the contrasting pair that is assimilated as uncategorizable speech sound or nonspeech sound. As some English vowels could be assimilated to two different Cantonese vowels based on the similarity of spatial proximity of constriction locations, those English vowel contrasts may have two possible pairwise assimilation patterns. According to the PAM, the gradient of discrimination levels should be in order, TC, CG and SC respectively, with TC being excellent, CG being fair to good and SC being poor.

Table 1. Assimilation patterns and its prediction of discriminability of English vowel contrasts.

Contrast	Assimilated Cantonese categories		Assimilation type	Prediction
/iː, I/	/iː/ -> [i]	/I/ -> [i] or [e]	SC/ TC	Poor/ Excellent
/iː, e/	/iː/ -> [i]	/e/ -> [e]	TC	Excellent
/iː, æ/	/iː/ -> [i]	/æ/ -> [e]	TC	Excellent
/I, e/	/I/ -> [i] or [e]	/e/ -> [e]	TC/ CG	Excellent/ Fair
/I, æ/	/I/ -> [i] or [e]	/æ/ -> [e]	TC/ CG	Excellent/ Good
/e, æ/	/e/ -> [e]	/æ/ -> [e]	CG	Fair
/uː, ʊ/	/uː/ -> [u]	/ʊ/ -> [u] or [ɔ]	SC/ TC	Poor/ Excellent
/uː, ɔː/	/uː/ -> [u]	/ɔː/ -> [ɔ]	TC	Excellent
/uː, ɐ/	/uː/ -> [u]	/ɐ/ -> [ɔ] or [aː]	TC	Excellent
/uː, aː/	/uː/ -> [u]	/aː/ -> [aː]	TC	Excellent
/ʊ, ɔː/	/ʊ/ -> [u] or [ɔ]	/ɔː/ -> [ɔ]	TC/ CG	Excellent/ Good
/ʊ, ɐ/	/ʊ/ -> [u] or [ɔ]	/ɐ/ -> [ɔ] or [aː]	TC/ CG	Excellent/ Good
/ʊ, aː/	/ʊ/ -> [u] or [ɔ]	/aː/ -> [aː]	TC	Excellent
/ɔː, ɐ/	/ɔː/ -> [ɔ]	/ɐ/ -> [ɔ] or [aː]	SC/ TC	Poor/ Excellent
/ɔː, aː/	/ɔː/ -> [ɔ]	/aː/ -> [aː]	TC	Excellent
/ɐ, aː/	/ɐ/ -> [ɔ] or [aː]	/aː/ -> [aː]	TC/ CG	Excellent/ Fair

Note. Assimilated Cantonese categories indicates the assimilation of English vowel to Cantonese vowel/s. TC represents Two-Category Assimilation. CS represents Category-Goodness Difference. SC represents Single-Category Assimilation.

Methodology

Test stimuli

This study tested the perception of four English front vowels (i.e., /iː/, /I/, /e/ and /æ/) and five English back vowels (i.e., /uː/, /ʊ/, /ɔː/, /ɐ/ and /aː/). Central vowels were excluded in this study to avoid confusion in the result as central and back vowels are frequently grouped together as a natural class (Chomsky & Halle, 1968; Kenstowicz, 1994). The test stimuli were English monosyllabic words in the form of either consonant-vowel or consonant-vowel-consonant structures. Thirty-two stimuli were selected and they were grouped into 16 pairs in the discrimination task and the identification task, including six pairs of front vowel stimuli and ten pairs of back vowel stimuli (See Appendix B for the list of stimuli). Words were in minimal pair relationship, which only differed in vowel. Front-back vowel contrasts were not tested as the acoustic differences such as F2 frequency, F2-F1 difference and F3-F1

difference are significant, correct perception would be relatively easy to achieve (Kent & Read, 2001).

All stimuli were recorded in a sound treated room with an industry-standard, highly versatile cardioid dynamic vocal microphone SHURE SM58 and a power amplifier M-AUDIO. An adult male native speaker of British English produced each stimulus thrice in isolation with an interstimulus interval of one second in a carrier phrase, 'Please point to __ __ __'. The recorded speech samples were then digitized and edited on computers using the audio recorder and editor software Audacity 2.0.5. The clearest and most stable production of each stimulus was then extracted from the carrier sentence. The same set of stimuli was used in all tasks.

The set of stimuli was spoken and visually presented in British English. There might be debates on the use of British English because television programmes, movies and commercials may not all be presented with a British accent. Also English teachers themselves may speak English with a mix of different accents or even an identifiable Hong Kong accent. Despite these, British English is the most widely accepted English language in teaching and learning as most of the textbooks are written in British English. Furthermore, Hong Kong was a British Dependent Territory under British administration from 1841 to 1997, British English is still highly preferred nowadays and most of the public facilities are spelled with British English. Therefore British English was used as the accent for the tasks.

Participants

Thirty-one Hong Kong native Cantonese speakers of EFL learners were recruited via convenient sampling from nine local mainstream primary schools to serve as unpaid participants. These included 15 females and 16 males aged from 8 to 11 at the time of the study (See Table 2 for the summary of participant background). Hearing screening, which consisted an otoscopic examination, an immittance audiometry and a pure-tone audiometry,

was conducted in a quiet environment to assess the participants' hearing ability. All the participants had passed the otoscopic examination with absent or non-occlusive amount of cerumen, the immittance audiometry with a tympanogram of type A and the pure-tone hearing screening at 25 dB HL at the octave frequencies of 500 to 4000 (American National Standard Institute, 1978; American Speech-Language-Hearing Association, 1985; Worrmald & Browning, 1996; Wong, Au & Wan, 2008). A short Chinese written questionnaire, which was translated from the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld & Kaushanskaya, 2007), was used to investigate the participants' language background. It was completed by the participants' caregiver. Participants were all Hong Kong native Cantonese speakers and had been resident in Hong Kong for at least six years. All of them started to learn English at the age of four years or younger. The participants reported that they had acquired their English mainly from school. Regarding the English exposure in English lessons at school, seven participants had been taught by native English-speaking teachers (NET), 18 of them had been taught by both NET and local teachers and six of them had only been taught by local teachers. According to the number of years of English education received and the quantity of native English exposure, this group of participants could be regarded as beginner level in English-language proficiency.

Table 2. Summary of participant background.

Gender (n)		Age (n)		Age		Years of English education		English exposure in English lessons (n)		
Male	Female	8;0 – 9;11	10;0 – 11;11	M	SD	M	SD	From NETs	From NET and local teachers	From local teachers
16	15	18	13	10.07	0.64	7.37	1.22	7	18	6

Note. M represents mean. SD represents standard deviation. NET represents native English speaking teacher.

Procedure

The research study was approved by the Faculty Research Ethics Committee of the University of Hong Kong. Informed written consent forms were obtained from the school principal and the participant's parent (See Appendix C for the school principal consent form) (See Appendix D for the parent consent form). Informed student assent forms were signed by each participants (See Appendix E for the student assent form).

After the participants had completed the language experience questionnaire and passed the pure-tone hearing screening, the following tasks were carried out. The tests were conducted in a quiet room where the participants were provided an iPad. All the stimuli were presented to the participants and binaurally over the iPad speaker at a comfortable listening level. A research assistant administered all the tasks and provided a short pre-task briefing together with task instructions in Cantonese.

Practice

The first practice was used to ensure the participants are familiar with the test stimuli. All 32 words were presented in pictorial and written forms on the iPad. The participants were allowed to click on each picture for as many times as they wanted to listen to the corresponding pronunciation. After the participants informed that they already had enough practice, the second practice was administrated to ensure the participants were familiar with all stimuli. The participants were presented with the 32 words, each with one trial. The target word was displayed on the iPad screen with two other words as distractors in a random order. All words were presented in pictorial and written forms. The participants were instructed to listen to the word and click the corresponding picture. A 100% accuracy was required before carrying out the discrimination and word identification tasks.

Discrimination task

The objective of this task was to examine how well the participants were able to discriminate the test stimuli presented in minimal word pairs. The ABX paradigm was used

and there were a total 32 trials as all 32 stimuli were tested. The target word (e.g., ship) was played through the iPad speaker with its minimal word pair in a random order (e.g., ship sheep sheep). In each trial, words were spoken in isolation with an interstimulus interval of 0.5 second. Three buttons with number 1, 2 and 3 were displayed on the iPad screen to indicate the words they had heard. As the word was spoken, the corresponding button will be enlarged. The participants were then asked to identify the odd word and to click the corresponding button.

Word identification task

The aim of this task was to investigate the participants' ability to identify a word from minimal word pairs. Thirty-two stimuli in 16 minimal pairs were all tested. Each word was targeted thrice, therefore each minimal pair was used six times. Thus, there were total of 96 trials. The target word (e.g., ship) was played through the iPad speaker and the pictorial and written forms of this minimal pair (e.g., ship sheep) were displayed on the iPad screen. The participants were then asked to identify the target word from the minimal pair and to click the corresponding picture.

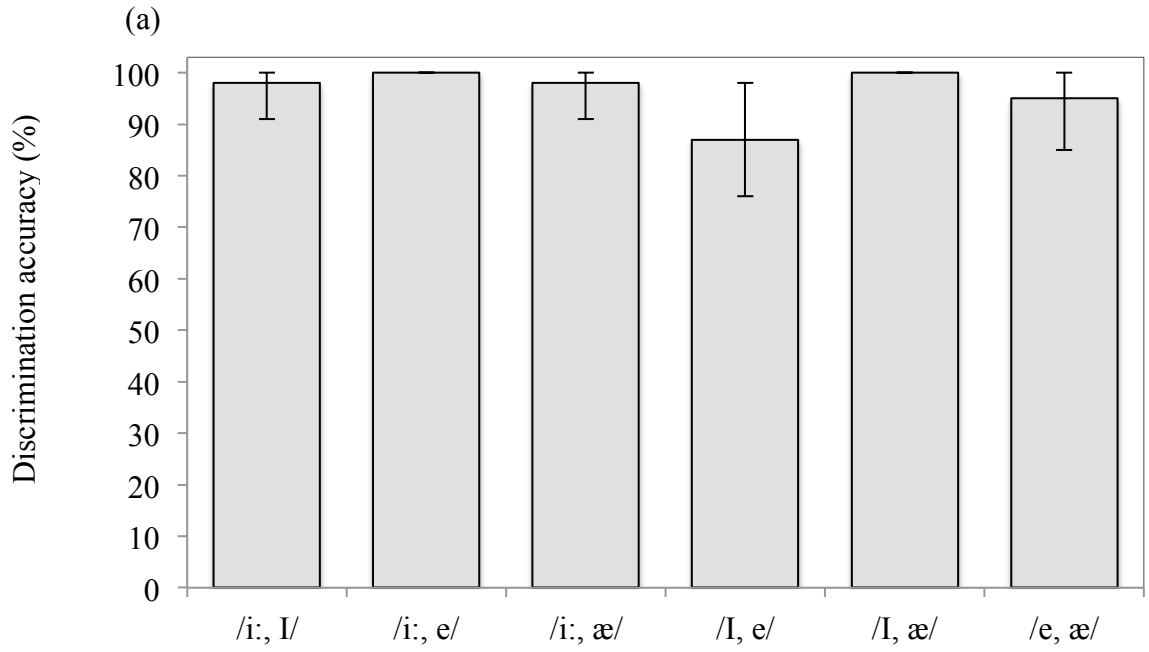
Result

Discrimination task

The significance of results from the ABX task was evaluated using the formula, $N/2 + \sqrt{N}$. If the number of correct trials exceeds $31/2 + \sqrt{32} = 22$, a significant result with 95% confidence was suggested. Eleven participants scored 32 as the highest mark and only two participants scored 28 as the lowest mark, with the rest of the participants scoring between these. As the numbers of correct trials from all participants exceeded 22, it can be concluded that the results were not due to chance.

A paired-samples t-test with a critical value of .01 showed that there was no significant difference between the overall mean discrimination accuracy of front vowels (M = 96%, SD = 4.84, SE = 2.00) and back vowels (M = 97%, SD = 3.13, SE = 1.28), $t(5) = -.11$, $p > .01$.

Figure 1 shows the participants' discrimination ability among different pairs of English vowels. The overall discrimination accuracy was 96% across all vowels. For English front vowels, the average accuracy was 96% (See Figure 1a). The participants' discrimination of the pairs /i:, e/ and /I, æ/ were the best, at 100% accuracy. The discrimination of the pair /I, e/, on the other hand, was the poorest, at 87% accuracy. For English back vowels, the average discrimination accuracy was 97% (See Figure 1b). Discrimination of the pairs /u:, ɔ:/, /u:, a;/, /ʊ, ɐ/ and / ʊ, a:/ were the best, at 100% accuracy. The discrimination of the pair /ɔ:, ɐ/ was the poorest, at 87% accuracy.



cont.

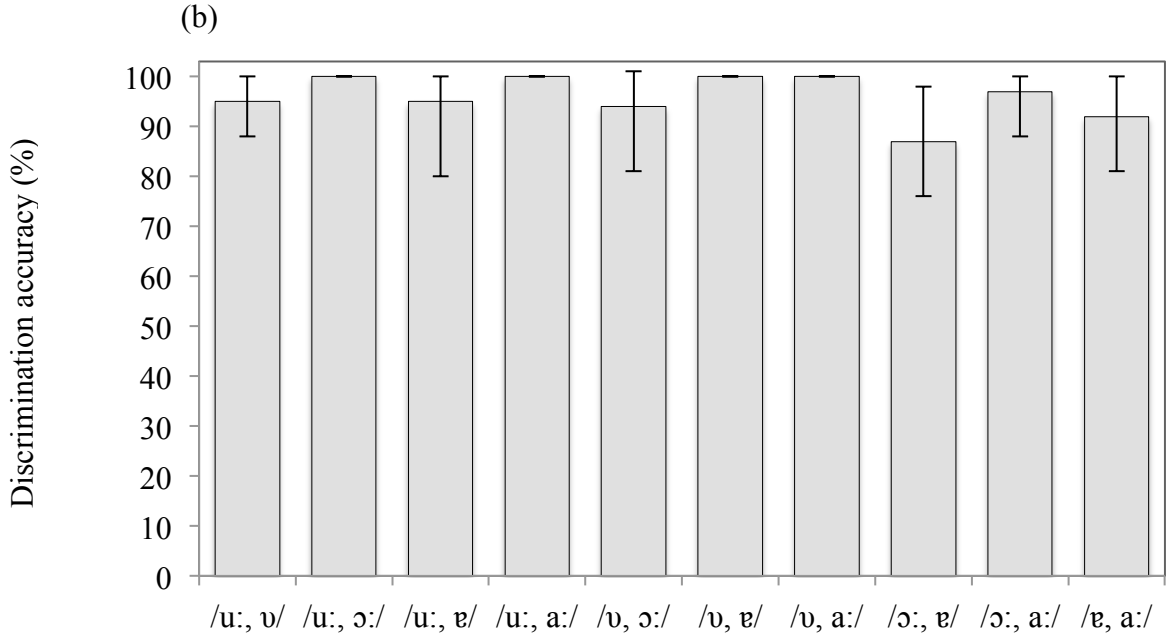


Figure 1. The mean percentage accuracy in discrimination between pairs of (a) front vowels and (b) back vowels, plotted with standard deviations as error bars.

Statistical analyses were then conducted using SPSS 22.0. To determine if there is any significant variance of discrimination ability among different pairs of English vowels, the correct discrimination scores of front vowel pairs and back vowel pairs were subjected to two one-way repeated measure ANOVAs separately. The within-subject factor was front vowel pairs (6 levels) versus back vowel pairs (10 levels) with significance set at .01 (2-tailed). There were significant main effects of vowel pairs in discrimination ability. For front vowel pairs, [$F(5, 150) = 5.45$, $p < .001$] and for back vowel pairs, [$F(9, 270) = 3.03$, $p < .005$]. However, pairwise comparisons with Bonferroni adjustment revealed no significant difference, between front vowel pairs and between back vowel pairs.

Word identification task

A paired-samples t-test with a critical value of .01 showed that the overall mean identification accuracy was significantly higher for back vowels (M = 86%, SD = 0.64 and SE = 0.01) than front vowels (M = 81%, SD = 0.96 and SE = 0.17), $t(30) = -2.88$, $p < .01$.

Figure 2 shows results of individual vowels from the identification task. For English front vowels, the identification of /æ/ was the best, at 87% accuracy. However the identification of /I/ was the poorest, at 73% accuracy. For English back vowels, the participants' identification of /a:/ was the best, at 94% accuracy. On the other hand, their identification of /ɐ/ was the poorest, at 75% accuracy.

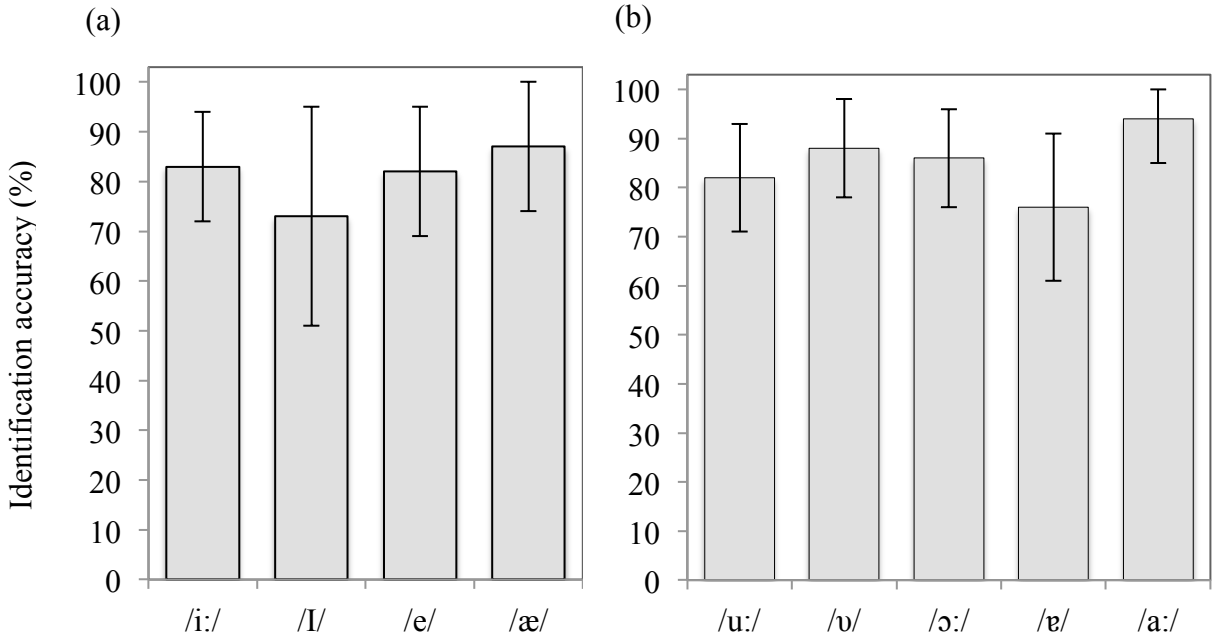


Figure 2. The mean percentage of identification accuracy among (a) front vowels and (b) back vowels. The error bars enclose the standard deviation for each vowel.

To determine if there is any significant variation of the identification ability among different English vowels, two one-way repeated measure ANOVAs with a critical value of .01 revealed a significant main effect of front vowels (4 levels), [$F(3, 90) = 5.02$, $p < .005$] and back vowels (5 levels), [$F(4, 120) = 13.56$, $p < .001$]. Pairwise comparisons with Bonferroni adjustment revealed no significant difference across front vowels. However, there was significant difference observed among the back vowels, when comparing the identification accuracy of /a:/ with /u:/, /ɔ:/ and /ɐ/ and /ɐ/ with /ɔ:/. Therefore it indicated that

/a:/ was significantly better than the rest of the back vowels whereas the identification accuracy of /ɐ/ was significantly lower than other back vowels.

As there were significant variances among different English vowels, one-way repeated measure ANOVAs and pairwise comparisons with significance set at .01 were conducted using SPSS 22.0 for each vowel to determine if there is any significant difference between different minimal pairs. Table 3 shows the participants' identification accuracy among different front vowel minimal pairs. The significant pairs are listed as follows, 1) for the minimal pairs that contain /i:/ as the target vowel, significant difference of the identification accuracy was found [$F(2, 60) = 55.32$, $p < .001$] and the accuracy of the pair /i:, I/ was significantly lower than the other two pairs; 2) for the minimal pairs of /I/, significant difference of the identification accuracy was also obtained [$F(2, 60) = 7.57$, $p < .005$] and the identification of the pair /I, æ/ was significantly better than other pairs; 3) for the minimal pairs of /e/, the difference of the identification accuracy was statistically significant [$F(2, 60) = 14.92$, $p < .001$] and the accuracy of the pair /e, æ/ was significantly lower than the rest; 4) for the minimal pairs of /æ/, significant difference of the identification accuracy was also found [$F(2, 60) = 11.15$, $p < .001$] and similar to the pairwise comparisons results of /e/, the accuracy of /æ, e/ was significantly lower than other pairs.

Table 3. A confusion matrix based on results from the identification task involving front vowel minimal pairs.

Target / Distractor	Identification accuracy (%)			
	/i:/	/I/	/e/	/æ/
/i:/		65 (32)	96 (11)	90 (19)
/I/	56 (27)***		88 (25)	95 (14)
/e/	96 (11)	67 (35)		71 (27)*
/æ/	97 (10)	89 (26)**	62 (32)**	

Note. The standard deviation of each pair is shown in bracket. The symbol, * indicates the pair was significantly worse or better than the other minimal pairs that contain the same target vowel, * for $p < .05$, ** for $p < .01$ and *** for $p < .001$.

Table 4 shows the participants' identification accuracy among different back vowel minimal pairs. The significant pairs are listed as follows, 1) for the minimal pairs involving /u:/, the identification accuracy was significantly different [$F(3, 90) = 18.78$, $p < .001$], the identification accuracy of /u:, ʊ/ and /u:, ɐ/ were significantly lower whereas the identification accuracy of /u:, ɔ:/ and /u:, a:/ were significantly higher; 2) for the minimal pairs of /ʊ/, significant difference of identification accuracy was also found [$F(3, 90) = 10.25$, $p < .001$] and same as the pairwise comparisons results of /u:/, the accuracy of /u:, ʊ/ was significantly lower than other pairs; 3) for the minimal pairs of /ɔ:/, significant difference of identification accuracy was also obtained [$F(3,90) = 14.01$, $p < .001$] and the accuracy of the pair /ɔ:, ɐ/ was significantly lower than the rest; 4) for the minimal pairs of /ɐ/, the difference of the identification accuracy was statistically significant [$F(3, 90) = 9.24$, $p < .001$] and the identification of the pair /ɐ, ʊ/ was significantly better than other pairs. Unlike other vowels, difference between the identification accuracy of the minimal pairs involving /a:/ was statistically non-significant [$F(3,90) = 3.11$, $p > .01$].

Table 4. A confusion matrix based on results from the identification task involving back vowel minimal pairs.

Target vowel / Distractor	Identification accuracy (%)				
	u:	ʊ	ɔ:	ɐ	a:
u:		69 (35)*	92 (21)	58 (41)	97 (13)
ʊ	59 (35)***		91 (23)	97 (10)**	95 (15)
ɔ:	99 (6)	94 (13)		71 (27)	98 (8)
ɐ	75 (29)**	96 (11)	63 (29)**		88 (20)
a:	96 (11)	94 (19)	97 (13)	73 (30)	

Note. The standard deviation of each pairs is shown in bracket. The symbol, * indicates the pair was significantly worse or better than the other minimal pairs that contain the same target vowel, * for $p < .05$, ** for $p < .01$ and *** for $p < .001$.

An independent factorial ANOVA was carried out to determine if gender, age and the quantity of exposure to native English influenced the participants' accuracy in identifying English vowels from minimal pairs. The overall identification accuracy was computed using

SPSS 22.0 in which gender (2 levels), age (2 levels) and quantity of exposure to native English (3 levels) served as between-subjects factors. In this analysis, there was no significant gender difference in identification accuracy (87% vs. 81%, $p > .01$). Similarly, age effects on identification accuracy was not statistically significant (83% in the age 8;0 – 9;11 group and 85% in the age 10;0 – 11;11 group, $p > .01$). However, there was a significant main effect of the quantity of native English exposure on the identification ability ($p < .001$). Participants who experienced English teaching from NET achieved a mean identification accuracy of 92%, participants who experienced English teaching from both NET and local teachers achieved a mean accuracy of 83% and participants who only experienced English teaching from local teachers achieved a mean accuracy of 77%. From the SPSS output, there was no interaction effect between quantity of native English exposure and gender as well as age. Multiple comparisons with Hochberg's GT2 also showed that the quantity of native English exposure had a significant positive effect on the identification accuracy.

Comparison between discrimination and word identification

Table 5 shows the categorization of English vowel contrasts according to PAM and their performance in both discrimination task and word identification task. The mean accuracy of discrimination task was compared to the mean accuracy of word identification task to determine if there is any significant difference in discrimination and word identification. A paired-samples t-test with significance set at .01 showed that results obtained on the discrimination task (M = 96%, SD = 4.4 and SE = 1.1) were statistically better than those on the word identification task (M = 84%, SD = 14.2 and SE = 3.6), ($p < .005$).

Table 5. Categorization of English vowel contrast as predicted based on the PAM and the actual performance in both discrimination task and word identification task.

Contrast	Assimilation -type	Prediction	Performance with mean accuracy (%)	
			Discrimination	Word - identification
/iː, I/	TC/ SC	Excellent/ Poor	Excellent (98) -	Poor (61) -
/iː, e/	TC	Excellent	Excellent (100) -	Excellent (96) -
/iː, æ/	TC	Excellent	Excellent (98) -	Excellent (94) -
/I, e/	TC/ CG	Excellent/ Fair	Good (87) ∨ / ∧	Fair (78) -
/I, æ/	TC/ CG	Excellent/ Good	Excellent (100) -	Excellent (92) -
/e, æ/	CG	Fair	Excellent (95) ∧	Poor (67) ∨
/uː, ʊ/	TC/ SC	Excellent/ Poor	Excellent (95) -	Poor (64) -
/uː, ɔː/	TC	Excellent	Excellent (100) -	Excellent (96) -
/uː, ɐ/	TC	Excellent	Excellent (95) -	Poor (67) ∨
/uː, aː/	TC	Excellent	Excellent (100) -	Excellent (97) -
/ʊ, ɔː/	TC/ CG	Excellent/ Good	Excellent (94) -	Excellent (93) -
/ʊ, ɐ/	TC/ CG	Excellent/ Good	Excellent (100) -	Excellent (97) -
/ʊ, aː/	TC	Excellent	Excellent (100) -	Excellent (95) -
/ɔː, ɐ/	TC/ SC	Excellent/ Poor	Good (87) ∨ / ∧	Poor (67) -
/ɔː, aː/	TC	Excellent	Excellent (97) -	Excellent (98) -
/ɐ, aː/	TC/ CG	Excellent/ Fair	Excellent (92) -	Good (81) ∨ / ∧

Note: The symbol, - indicates that the performance matched with the PAM prediction, the symbol, ∧ indicates that the performance was better than the prediction and the symbol, ∨ indicates that the performance was worst than the prediction.

Discussion

Discrimination

In this section, the results of the discrimination task are discussed. The results are largely consistent with predictions made based on the PAM that the discrimination level of TC was better than CG and CG was better than SC. Based on the PAM, predictions and actual performance of the ability to discriminate vowels were made and listed in Table 5. For front vowels, participants' discrimination of /iː, e/ and /I, æ/ (100% accuracy) were the best whereas /I, e/ (87% accuracy) was the poorest. As mentioned above, some of the English vowel contrasts can have two possible pairwise assimilation patterns, the high accuracy of /iː,

I/ and /I, æ/ (98% and 100% accuracy) may suggest that these pairs of contrasts were perceived as TC type during the discrimination, however, the relatively low discrimination accuracy of /I, e/ (87% accuracy) suggests that it is possibly categorized as the CG type rather than TC type. In order to avoid the uncertainty of assimilation type and achieve more precise predictions, perceptual goodness rating is suggested in future research.

For back vowels, The participants' discrimination of /u:, ɔ:/, /u:, a;/, /ʊ, ɐ/ and / ʊ, a:/ (100% accuracy) were the best and these vowel contrasts are all categorized as TC type. The high accuracy of /u:, ʊ/, /ʊ, ɔ:/and /ʊ, ɐ/ (scores at or above 94) suggests that these contrasts were also being perceived as TC type in discrimination. The relatively low discrimination accuracy of /ɐ, a:/ (92% accuracy) suggests that it is possibly categorized as the CG type rather than TC type and /ɔ:, ɐ/ was most poorly discriminated (87% accuracy), therefore this contrast was probably being categorized as SC type.

Word identification

A significant difference was found between the identification accuracy of front vowel minimal pairs and back vowel minimal pairs. As back vowels had a higher mean score, it is concluded that back vowels achieve better performance than front vowels. According to Kent and Read (2001), front vowels and back vowels differ in some acoustic features which include F2 frequency, F2-F1 difference and F3-F2 difference. Therefore, a combination of these factors leads to a better performance of back vowels. As acoustic measurements were not involved in this study, it can not be specified which feature/s is dominated in resulting a better perception of back vowels than front vowels, further research with measurements of formant frequencies is needed to draw a clearer conclusion.

To view the English vowels individually rather than in a minimal vowel contrast, the overall accuracy of /a:/ was the best and /I/ and /ɐ/ were comparatively poor. The high accuracy rate of the English vowel /a:/ may be due to the high similarity to the Cantonese /a:/.

The English vowel /a:/ is reported as having similar articulation as the Cantonese vowel /a:/ and both of them are low back vowels (Chan, 1968; Chan & Li, 2000). For the low accuracy rate of the English vowels /I/ and /ɐ/, it maybe due to the close position for the English vowel /I/ to the Cantonese vowels /i/ and /e/ and for the English vowel /ɐ/ to the Cantonese vowel /ɔ:/ and /a:/, so they assimilated to both Cantonese vowels respectively. Low identification accuracies in the contrasts /i:, I/ and /I, e/ as well as /ɐ, ɔ:/ and /ɐ, ɔ:/, contributed to the low overall accuracy of /I/ and /ɐ/.

The results of the English front vowel contrasts in the word identification task were highly consistent with predictions made based on the PAM. The TC type, /i:, e/, /i:, æ/ and /I, æ/ had an excellent identification accuracy whereas the CG type /I, e/ had a fair identification accuracy and the CG type /e, æ/ and the SC type /i:, I/ had poor identification accuracy. The low identification accuracy of /i:, I/ and /I, e/ was possibly due to the assimilation of /I/ to the Cantonese vowels /i/ and /e/. The English vowels /i:/ and /I/ are both acceptable exemplars of the Cantonese vowel /i/, and the English vowels /I/ and /e/ are exemplars of the Cantonese vowel /e/ with fairly similar magnitude of difference in category goodness as they have rather similar articulation, except the latter is less close than the former. Therefore, the identification of these contrasts was poor. For the low identification accuracy of /e, æ/, it could be explained that the English vowel /æ/ does not have any counterpart in Cantonese and there is no Cantonese vowel in the low front position, therefore /æ/ is perceived as the English vowel /e/ and they are both assimilated to the Cantonese vowel /e/.

Similar to the results of the English front vowel contrasts, the results of the English back vowel contrasts in word identification were also consistent to a great extent with predictions made based on the PAM. The TC type vowel contrasts (i.e., /u:, ɔ:/, /u:, a:/, /ʊ, ɔ:/, /ʊ, ɐ/, /ʊ, a:/ and /ɔ:, a:/) had an excellent identification accuracy. The CG type vowel contrast, /ɐ, a:/ yielded a good identification accuracy and the SC type vowel contrasts, /u:, ʊ/ and /ɔ:, ɐ/

resulted in relatively poorer accuracy. It is not unexpected for the SC type, /u:, ʊ/ and /ɔ:, ɐ/ to yield poor identification as both vowels in the contrasts are acceptable exemplars of the Cantonese vowel /u/ and /ɔ/ correspondingly. However it is surprisingly to note that the English vowel contrast, /u:, ɐ/, which at the outset of the study was expected to be of the TC type yielded poor identification accuracy. Due to low proficiency of the participants and limited number of contrastive words that could be illustrated, the two vowels were represented using the words 'Boom' and 'Bomb' while these word stimuli share common semantic meaning and the word 'Boom' was unfamiliar to some of the participants. These factors might have caused lower identification accuracy.

Results from the independent factorial ANOVA showed that the quantity of English exposure was a significant factor related to the accuracy of vowel identification. Best and Tyler (2007) also found that language exposure such as quantity of input from native English speakers plays an important role in language acquisition. Learning English from native English speakers could improve perception of EFL listeners as their English phonetic inventory can be also elaborated (Flege, Bohn & Jang, 1997; Best & Tyler, 2007). Flege (1995) explained that language experience of a non-native language over time allows learners to notice some cross-language phonetic differences and therefore they could ultimately establish a new phonetic category to represent phonemes in that non-native language.

Comparison between discrimination and word identification

Significant difference was found between the performances in discrimination task and word identification task. Although the results of both discrimination task and word identification task follow the assimilation predictions of the PAM which explains the perception of EFL learners on the basis of articulatory-phonetic similarities between Cantonese and English vowels, the contrasts were perceived more accurately in the discrimination task than the word identification task. The cognitive neuropsychology model

explains that higher cognitive functional load is required for the word identification task than the discrimination task (Ellis & Young, 1988). For discrimination, the participants have to detect the differences of the auditory features such as formant frequencies in the minimal vowel pair (Ryalls, 1996). But for word identification, in addition to the discrimination of auditory features, cognitive processes such as phonological lexicon, semantic system and visual object recognition system as well as the evaluation of information against long-term memory are also necessary (Ellis & Young, 1988; Flege, 2003). Therefore it is not surprising that word identification was a bit more difficult to these EFL learners.

According to results listed in Table 5, performance on the discrimination task was better than that on the word identification task for some minimal vowel contrasts, namely /i:, I/, /I, e/, /e, æ/, /u:, ʊ/, /u:, ɐ/, /ɔ:, ɐ/ and /ɐ, a:/. A common acoustic feature is observed between these vowel contrasts is the tenseness; that is, one of the vowels in these contrasts was a tense vowel (i.e., /i:/, /e/, /u:/, /ɔ:/ and /a:/) and the contrastive partner was a lax vowel (i.e., /I/, /æ/, /ʊ/ and /ɐ/). Tense-lax feature of a vowel is one of the major factors that influence vowel duration. Kent and Read (2001) inform that other than formant frequencies and formant pattern, vowel duration is an important parameter for the acoustic specification of vowels. Therefore apart from the difference in difficulty level of listening tasks (i.e., discrimination vs. word identification), vowel duration may also be one of the factors that contributes to the performance difference in these two tasks. In previous studies, EFL listeners were found to over-rely on vowel duration, that is they either used vowel duration exclusively or they used more than one cue but still weighted vowel duration as the primary perceptual cue to signal the vowel contrast (Flege et al., 1997; Escudero, 2000). In the current study, the participants were able to use durational cues in discriminating vowels but were not able to use the same cues in an identification task where only the target vowel was presented. Yet, Flege (2003) suggested that a better identification task should involve more than two choices in order to

reduce the chance level and also provide evaluation of the use of frequency and temporal cues. The importance of durational cues could be evaluated in future research.

It is also noticed that the results of the discrimination task are considerably better than predictions made based on the PAM whereas the results of the identification task are highly consistent with the predictions. The better performance in discrimination than the PAM assimilation predictions can be explained by the effects of experience on English. Although the participants are categorized as beginning learners of EFL, the language background questionnaire revealed that they have all been learning English for at least five years. As suggested by Flege, Bohn and Jang (1997), the perception accuracy of English vowels improves with English language experience. In addition, the word identification task, as mentioned above, is more difficult than the discrimination task. Ceiling effects were noted in both tasks but were particularly greater in the discrimination task. As the testing in the current study involved testing in quiet only, it is recommended that further testing can be conducted in noise in order to tax the auditory system and reduce the influences from ceiling effects.

Theoretical and pedagogical implications

Results from the current research have revealed both theoretical and pedagogical significance. Theoretically, the results provide evidence that is highly consistent with the PAM, emphasizing the effects of similarity of spatial proximity of constriction locations between non-native vowels and native vowels and confirming the applicability of the PAM in predicting English vowel perception by native Cantonese speakers who are beginning EFL learners.

Other findings of the present study that do not relate to the PAM are also enlightening. The differences between the discrimination task results and the word identification task result as well as the performance difference between front and back vowels both suggest that

acoustic features such as formant frequencies and vowel duration also play an important role in foreign vowel perception. Therefore, future study about English vowel perception of EFL learners is suggested to take these parameters into account.

Pedagogically, the results of the word identification task highlight a perception problem. The Hong Kong native Cantonese EFL learners encounter difficulties when identifying particular English vowel pairs, namely /i:, I/, /I, e/, /e, æ/, /u:, ʊ/, /u:, ɐ/ and /ɔ:, ɐ/. English teachers are recommended to focus students' attention on these problematic and confusable English sound pairs and assist them in identifying between them.

Results of the identification task also have pedagogical significance. The results provide evidence that participants with greater quantity of native English exposure in English lesson at school achieved much higher word identification accuracy. It is suggested that receiving English teaching from NETs can improve the perception accuracy of English vowels. Therefore, schools are advised to arrange some English lesson from NET for their students in order to provide more native English input. The quantity requires to make a significant difference should be examined in future studies.

Limitations

Some limitations in the current study are noted. First, only English vowels are investigated. Therefore, the results can only be applied on the perception of vowels but not consonants.

Secondly, only beginning learners of English were involved. Therefore, the results can not be generalized to all Cantonese English learners at different English proficiency levels. Furthermore, it can not show the age-related developmental tendency of English vowel perception by Cantonese English learners.

Thirdly, the choice of stimuli also contributed to the limitations. As some of the stimuli such as 'Boom', were unfamiliar to the participants, the participants may identify those

vowels upon memory form the practice, the effects of familiarity with the test stimuli are unknown. Moreover, some of the stimuli shared similar semantic features such as 'Boom' and 'Bomb', confusion might have occurred when the participants were asked to choose the corresponding pictorial form of the target.

An inadequacy is concerned with the perceptual goodness difference between English and Cantonese vowels. According to Best (1995), "Assimilation is assumed to be tapped by tests that measure identification (labeling), classification, or categorization (including goodness ratings) of non-native phones" (p. 194). As the above tests have not been done in this study, the assimilation type and its prediction were uncertain and imprecise as it was only depended on the overview of English and Cantonese vowels. By adding perceptual goodness rating in future studies, degrees of similarity between English and Cantonese vowels perceived by the listeners can be clearly illustrated. As suggested by Strange et al (1998) and Chan (2012), it can be carried out by first having the participants listen to an English vowel, and then classify the English vowel as one of the Cantonese vowels, after that the participants have to rate the English vowel for the degree of similarity to the Cantonese vowel by using an interval scale.

Another inadequacy is the lack of acoustic measurement. In this study, no acoustic measurement of formant frequencies, vowel duration and spectral patterns has been carried out, therefore conclusions could only be made based on results from the current study and previous studies (Chan, 1968; Chan & Li, 2000, Kent & Read, 2001; Chan, 2012). Therefore acoustic measurements of the target vowels are recommended to be carried out.

Conclusion

In this research paper, I have described the results of a research study that investigated the perception of English vowels by Hong Kong native Cantonese EFL primary school students. The applicability of the PAM to the vowel perception by beginning learners was

also examined. It is found that the results of the English vowel discrimination and identification are greatly consistent with predictions made based on the PAM.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Lena Wong. She provided me with very valuable advice and guidance throughout the planning and development of this study. Technical assistance provided by Dr. Felix Chen was also greatly appreciated. I would also like to thank the participants and the following organization and school for their assistance with the data collection: Mentality & Art Association and S.K.H. Lee Shui Keung Primary School. At last, my special thanks are extended to my family and friends for the unconditional support.

References

American National Standard Institute. (1978). *Methods for manual pure-tone threshold audiometry.* New York: ANSI.

American Speech-Language-Hearing Association. (1985). Guidelines for identification audiometry. *ASHA, 27,* 49-53.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phoenetic dissimilarity and L2 speech learning the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32,* 233-250.

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. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language research* (pp. 171–204). Baltimore: York 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*(2), 775-794.

Best, C. T., & Tyler, M. D. (2006). Nonnative and second-language speech perception: commonalities and complementarities. In M. J. Munro & O.-S (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Bohn, O-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11,* 303-328.

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.

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.

Chan, Y-Y. (1968). *The spectral properties of Cantonese vowels: comparison with English vowels.* Pokfulam, HK: The University of Hong Kong.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New york: Harper and Row.

Ellis, a. W., & Young, A. W. (1988). *Human Cognitive Neuropsychology.* Hove, UK: Erlbaum.

Escudero, P. (2000). *Developmental patterns in the adult L2 acquisition of new contrasts: the acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers.* University of Edinburgh, Scotland, UK.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(4)*,* 551-585.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15,* 47-65.

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. 229–273). Baltimore: York Press.

Flege, J. E., Bohn, O-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25,* 437-470.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production.* Berlin: Mouton de Gruyter.

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America, 97*(4), 2540-2551.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*(5), 2711-2724.

Ingram, J. C. L., & Park, S-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25,* 343-370.

Kenstowicz, M. (1994). *Phonology in generative grammar.* Cambridge, Mass: Blackwell.

Kent, R. D., & Read, C. (2001). *Acoustic Analysis of Speech*. (2nd ed.). San Diego: Singular Publishing Group.

Marin, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profile in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50*(4), 940-967.

Pilus, Z. (2005). Perception of voicing in English word-final obstruents by Malay speakers of English: Examining the Perceptual Assimilation Model. *Malaysian Journal of ELT Research, Inaugural Volume,* 1-12.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America, 97*(2), 1286-1296.

Ryalls, J. (1996). *A basic introduction to speech perception.* San Diego, CA: Singular Publishing Group.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent. S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics, 26,* 311-344.

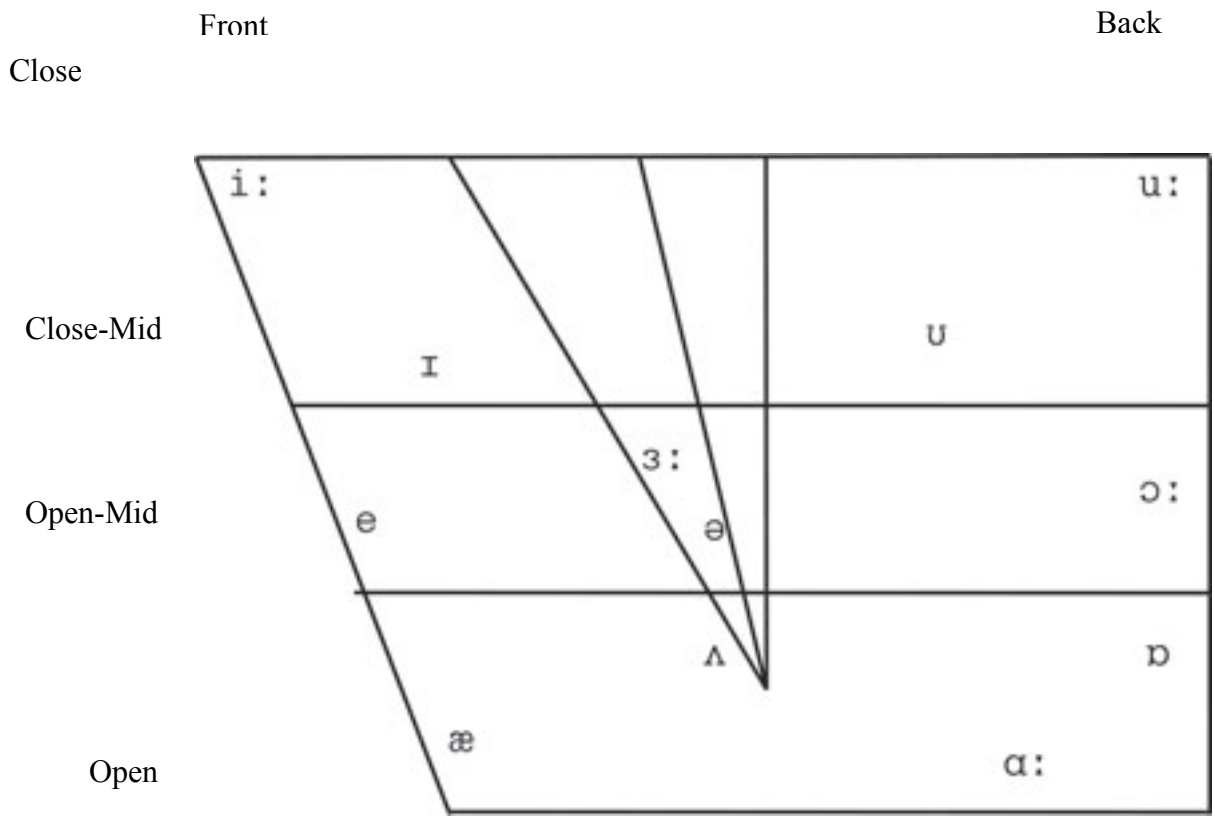
Wong, L. L. N., Au, J. W. Y., & Wan, I. K. K. (2008). Tympanometric characteristics in Chinese school-aged children. *Ear & Hearing, 29*(2), 158-168.

Worrmald, P. J., & Browning, G. G. (1996). *Otoscopy: a structured approach.* San Diego, CA: Singular Publishing Group.

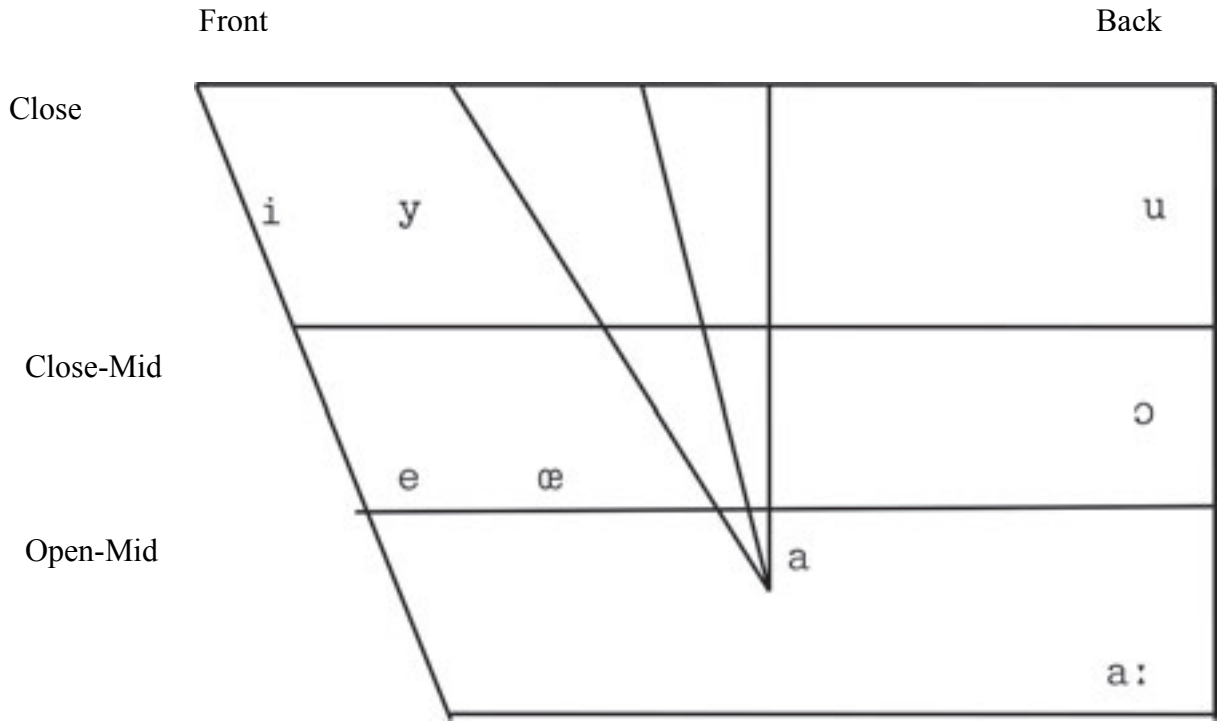
Appendix A

Overview of English and Cantonese vowels (Chan & Li, 2000)

Overview of English vowels



Overview of Cantonese vowels



Appendix B

List of stimuli

Minimal word pairs for front vowels		
1	/iː, I/	Sheep vs. Ship
2	/iː, e/	Read vs. Red
3	/iː, æ/	Seed vs. Sad
4	/I, e/	Pin vs. Pen
5	/I, æ/	Hit vs. Hat
6	/e, æ/	Bed vs. Bad

Minimal word pairs for back vowels		
1	/uː, ʊ/	Pool vs. Pull
2	/uː, ɔː/	Cool vs. Call
3	/uː, ɐ/	Boom vs. Bomb
4	/uː, aː/	Juice vs. Jars
5	/ʊ, ɔː/	Full vs. Fall
6	/ʊ, ɐ/	Good vs. God
7	/ʊ, aː/	Book vs. Bark
8	/ɔː, ɐ/	Forks vs. Fox
9	/ɔː, aː/	Four vs. Far
10	/ɐ, aː/	Hot vs. Heart

Appendix C

Principal consent form

THE UNIVERSITY OF HONG KONG
Faculty of Education

8th November 2013

Dear Principal,

I am Pang Hiu Wa, a year 4 Speech and Hearing Sciences student from the Faculty of Education at University of Hong Kong. As part of my bachelor degree, I am required to conduct a small-scale research on 'The perception of English vowels by native Cantonese English as a second language (ESL) primary school students'. Acquisition of English is always one of the most important issues for students in Hong Kong. However English learning can be challenging for students as ESL learners usually have demonstrated difficulties in perceiving certain English speech sounds. In the present research, the perception of English vowels will be examined in order to help investigate the English learning situation in Hong Kong. Findings from this research can contribute to the understanding of ESL learners with beginner English proficiency level and perhaps in the future the information can be used to assist them to acquire English with greater ease.

I would like to invite around 25 students from grades 4 to 6 from your school to participate and all the participants will be (1) native speakers of Cantonese; (2) learning English as a foreign language; and (3) having normal language development with no reported learning difficulties or other disabilities. Moreover, they should not have resided in an English-speaking environment and/or received any form of phonetics training prior to the research.

According to the University's policy on the ethical conduct of research, I am writing to seek your consent for recruiting students for this research.

This research will involve a demographic questionnaire to investigate the participants' language background, a pure tone audiometry to assess their hearing ability and a few tests to identify the English words being heard. The questionnaire will be distributed together with the parent's consent form. The school will then collect it back together with the reply slip before test day. It will take around 10 minutes for the parents of participants to complete the questionnaire. The other tests will be carried out either in a quiet classroom at your school or in a sound booth at Meng Wah Building at the University of Hong Kong according to the parent's preference. All test stimuli will be presented over headphones at a comfortable listening level. Each step will last for around 10 minutes. By estimation, each participant will be required to spend 60 minutes to complete all the test procedures and a short break will be given in between tests.

There is no known risk associated with this study. The information collected will only be used for the aforementioned study only. Strict confidentiality will be maintained. The questionnaire will be stored in a filing cabinet and the data will be recorded and stored in a password-protected computer program in a password-protected laptop computer that is only accessible to the researcher and supervisor during the research period, now till end of June in

2014. Students' participation is voluntary and participants can withdraw from the study at any time.

If concerns arise about this aspect of my work, please feel free to contact me at 6092-0445/ hwpang16@hku.hk, or my supervisor Dr Lena L N Wong at 2859-0590/ llnwong@hku.hk. If you have questions about your students' rights as a research participant, please contact the Human Research Ethics Committee for Non-Clinical Faculties, HKU at 2241-5267.

If you agree to these procedures, please sign one copy of this letter and return it to me.

Thank you for your attention and support. Your help is very much appreciated.

Yours Sincerely,

Pang Hiu Wa
Division of Speech and Hearing Sciences
Faculty of Education
The University of Hong Kong

I agree to the procedures set out above to facilitate the following students to conduct the research project in my school.

Endorsed by: Date:

Appendix D

Parent consent form

父母/監護人同意書

敬啓者：

本人彭曉華是香港大學教育學院言語及聽覺科學系四年級生。本人將會進行一項關於英語韻母感知的研究， 對象為以廣東話為母語及以英語為第二語言的小學生， 希望能夠邀請貴子女參加。 在現今的香港， 英語學習對學生而言是非常重要的，但與此同時它對於一些學生而言亦充滿了挑戰性。本研究旨在探討學生的英語韻母感知，從而了解香港的英語學習情況。 研究結果可以幫助我們理解以英語為第二語言的英語初學者， 而研究資料有望能用於協助他們學習英語， 使英語學習變得較為容易。

一份有關語言背景的問卷會連同此同意書一同派發。 參與是次研究學生的家長將需要大約 10 分鐘完成這份問卷。而參與是次研究的學生將需要通過聽力審查 和完成一些有關識別英文單詞的測試。根據家長的選擇， 測試將於貴子女就讀學校的班房或香港大學明華綜合大樓的隔音房進行。 所有被測試的英文單詞將通過耳機於舒適的收聽音量播放。 每個步驟將持續約 10 分鐘。 每位參與者將需要大約 60 分鐘完成所有的測試程序， 測試與測試中間將會提供一個短暫的休息。

是次研究沒有任何已知的風險。 而完成所有的測試程序後， 你將會獲得一份簡短的口頭聽力報告。

所收集的資料將只作研究用途， 絕對保密。由現在起至二零一四年六月尾其間，研究所搜集的問卷將被存儲於文件櫃內而所搜集的數據將被記錄並存儲在受密碼保護的手提電腦及電腦程式中， 只有研究員和研究導師能夠讀取。 參與純屬自願性質，學生可以隨時退出研究。 如閣下對是項研究有任何查詢， 請隨時與本人聯絡 (6092-0445/ abbiepang@hotmail.com) 或與研究導師黃麗娜博士聯絡 (2859-0590/ llnwong@hku.hk)。如閣下想知道更多有關研究參與者的權益，請聯絡香港大學非臨床研究操守委員會(2241-5267)。

請閣下填妥以下回條及問卷並交回學校，以表示你是否同意貴子女參與是項研究。本人會盡快與你聯絡商討有關的細節。

非常感謝你的支持和幫助。

此致

香港大學教育學院言語及聽覺科學系

彭曉華謹啟

二零一三年十一月八日

家 長 回 條

學生姓名：_____ 班別：_____ 學號：_____

本人 ** 同意 / 不同意 子弟參與是項研究。
(**請刪去不適用者)

家長姓名：_____

家長簽署：_____

家長聯絡電話：_____

家長聯絡電郵：_____

日期：_____

Appendix E

Student assent form

Student Assent Form

THE UNIVERSITY OF HONG KONG

The perception of English vowels by native Cantonese English as a second language primary school students

Dear Students,

I am Pang Hiu Wa, a year 4 Speech and Hearing Sciences student from the Faculty of Education at the University of Hong Kong. My supervisor Dr Lena L N Wong and I are now conducting a study titled "The perception of English vowels by native Cantonese English as a second language primary school students". I hope to spend some time with you to let you know more about details of this project.

I have obtained written consent from your parent/guardian earlier to let you join this project. However, your decision is also very important to us. If you agree to join this project, you will be invited to complete a hearing screening, a spoken word-picture and written word matching task, an auditory word discrimination task and a word identification task. The tests will be carried out either at your school or at the University of Hong Kong according to your parents' preference. It will take around 60 minutes of your time to complete all the test procedures.

The data will be recorded and stored in a password-protected computer program in a password-protected laptop computer that is only accessible to my supervisor and me during the research period, now till end of June in 2014. If you have any further question, please raise it now. Thank you for your support.

If you agree to take part in this project, please put a tick in the following box and sign your name besides it.

☐ I agree to participate in this project. Signature:_____

OR

If you do not agree to take part in this project, please put a tick in the following box and sign your name besides it.

☐ I do not agree to participate in this project. Signature:_____

Student Name:_____ Class: _____ Date:_____

Yours sincerely,

Pang Hiu Wa
Division of Speech and Hearing Sciences
Faculty of Education
The University of Hong Kong