Title	Perception of English vowels as a foreign language of Hong Kong Cantonese secondary school speakers
Author(s)	Ng, Shuk-ki; 吳淑琪
Citation	Ng, S. [吳淑琪]. (2014). Perception of English vowels as a foreign language of Hong Kong Cantonese secondary school speakers. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2014
URL	http://hdl.handle.net/10722/238932
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Perception of English Vowels as a Foreign Language

of Hong Kong Cantonese Secondary School Speakers

Ng Shuk Ki

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

Perception of English Vowels as a Foreign Language of Hong Kong Cantonese Secondary School Speakers

Ng Shuk Ki

Abstract

This dissertation reports on the results of a research study that investigated the perception of English vowels by native Cantonese speakers who are learning English as a foreign language (EFL) as well as the applicability of the predictions of the Perception Assimilation Model (PAM) to foreign language perception by Hong Kong Cantonese EFL learners. Thirty-one local secondary school students participated in a perception test to discriminate and identify English minimal pairs. The results affirm the hypothesis of the PAM on the perceptual identification of different types of minimal pairs. The results also call for the usage of explicit English phonological training in Hong Kong Education to facilitate English acquisition.

Introduction

Literature Review

Foreign language learners often have difficulty in perceiving some vowels in the foreign language. This difficulty may be attributed to the fact that some vowels in the foreign language are different from those in learners' native language, that is, the vowel inventories in their native language and the foreign language are not exactly the same. In order to explain this phenomenon, speech perception models such as the Perceptual Assimilation Model (PAM) (Best, 1994) and the Speech Learning Model (SLM) (Flege, 1995) were proposed. Both of them relate the listeners' performance to perceive non-native phonemes to the phonemes in their native language (Kingston, 2003).

According to the PAM (Best, 1994), listeners perceptually assimilate non-native phonemes into their own phonemic inventory. It proposes three possible classifications according to the similarity of the gestural information between the phonemes in the non-native language and those in the native language. The first one is categorized exemplar of native phoneme. Obviously, this applies to the phonemes in the foreign language which can be clearly and easily assimilated to a particular native segmental category. The second one is uncategorized speech sound that falls in between specific native phonemes. The third one is the speech sounds which bear no detectable similarity to any native phonemes; they are not assimilated to speech and heard as non-speech sound. Based on the finding on the

perception of English vowels of Cantonese English-as-a-foreign-language (EFL) speakers by Chan (2012), all the English vowels (in the study) can be assimilated into specific Cantonese vowels, which is the first type of non-native phoneme assimilation in the PAM.

It should be emphasized that the perceptual accuracy of a single phoneme in foreign language could not be distinctively predicted by the PAM. Perception of a non-native phoneme is excellent does not necessarily mean perception of another non-native phoneme of the same type is excellent. For example, the perception of the non-native speech sounds of the first type, categorized exemplar of native phoneme, may range from excellent to poor. However, the assimilation of each phoneme in a contrast makes up pairwise assimilation patterns and these in turn predicts the degree of perception for diverse non-native contrasts (Best, 1995). There are six pairwise assimilation patterns. However, all the English vowels can be assimilated to specific Cantonese categories (Chan, 2012). As a result, there are only three possible pairwise assimilation patterns (Best, 1995):

1.) Two-Category Assimilation (TC type): Each English vowel is assimilated to two different Cantonese categories, excellent perception is expected.

2.) Category-Goodness Assimilation (CG type): Both English vowels are assimilated to the same Cantonese category but vary in the discrepancy from the native “ideal”, moderate to good perception is expected.

3.) Single-Category Assimilation (SC type): Both English sounds are assimilated to the same Cantonese category with the same magnitude of discrepancy from the native "ideal", perception is expected to be the poorest.

A study by Best, McRoberts, and Goodell (2001) supported the above-mentioned claims of the PAM by assessing native English speakers' perception of Zulu and Tigrinya contrasts. Results showed that perception of the TC type non-native contrasts was highly accurate; that of the CG type non-native contrasts was quite good and poor perception was found for the SC type non-native contrast. However, there is a lack of information affirmed the application of the PAM on Cantonese EFL listeners.

According to the SLM, perception of phonemes in second language (L2) is also linked to the similarity to phonemes in the first language (L1) (Flege, 1995). The SLM predicts that L2 learners will judge the L2 phonemes according to the existing phonemes of the L1 before establishing the L2 phonemic category. This will filter out the distinctive features of the L2 phonemes which are important phonetically. If a phonetic contrast differentiate only phonemes in L2 but not in L1, that is, there is no such a contrast on L1 (like the absence of tense-lax vowel contrast in Cantonese), it will be difficult for the L2 listener to detect and differentiate the L2 phonemes (Guion, Flege, Akahane-Yamada & Pruitt, 2000). For the L2 phonemes having similar L1 counterparts, as the distinctive features are filtered out, a new phonetic category is difficult to be formed. On the contrary, if the difference between an L2

phoneme and the L1 phoneme is larger, a new phonemic category could be established with ease so the perception of these phonemes is expected to be better.

Many studies have testified the applicability of the SLM. Aoyama, Flege, Guion, Akahan-Yamada, and Yamada (2004) investigated the perception of English consonants /r/ and /l/ in inexperienced Japanese English-as-a second-language (ESL) children. They found that English consonant /r/ which was more dissimilar to Japanese consonant/r/ than English consonant /l/ was better perceived. Chan (2012) also found that perception of English vowels by advanced Cantonese ESL speakers is highly related to the perceived similarity of the English phonemes to the Cantonese counterparts. The English phonemes which counterparts in Cantonese are more accurately identified than those do not have Cantonese counterparts. The accuracy of perception of other English phonemes depends on the degree of phonetic distance from closest Cantonese vowels. In Chan's study, only advanced English learners, that is, university students who major in English, were recruited as subjects. The results may not be generalized to populations of other age populations or with lower English proficiency.

Objective

It was evident that the PAM focused on explaining the perception of naïve nonnative learners while the SLM is more applicable to the advanced L2 speakers (Best & Tyler, 2006). It seems more suitable that to examine whether the PAM applies to the perception of English vowels by participants who exhibit low-to-intermediate English proficiency.

This study investigated the perception of English vowels by Hong Kong Cantonese-speaking local secondary School students to testify the applicability of the claims of the PAM for Cantonese EFL listeners. Perceptual discrimination of the English stimuli and perceptual identification between front vowels and between back vowels in the stimuli would be investigated. Comparison of perception among minimal pairs would be performed to assess perception of the stimuli pairs to see if the claims and prediction of the PAM apply.

For front vowels, according to a perceptual similarity test conducted by Chan (2012), most of the Cantonese ESL speaking subjects assimilated English vowels /iː/ and /ɪ/ into Cantonese vowel /i/. English vowels /e/ and /æ/ were assimilated to Cantonese vowel /e/. For back vowels, most of the Cantonese ESL subjects in Chan's study assimilated English vowels /uː/ and /ʊ/ into Cantonese vowel /u/. English vowels /ɔː/ and /ɒ/ were assimilated to Cantonese vowel /ɔ/, and English vowel /ɑː/ was not included in the study. However, with reference to the overview of English vowels and Cantonese vowels (Chan & Li, 2000), English vowel /ɑː/ was the most similar to Cantonese vowel /aː/ regarding the backness and height of the vowels. As the other English vowels are found to be assimilated to the Cantonese vowels with the most similar backness and height, assimilation of English vowel /ɑː/ to Cantonese vowel /aː/ will be used in this study.

With the above assimilation patterns of English vowels, we can conclude that all the minimal pairs in this study are of TC type in the PAM except the minimal pairs with English

vowel contrasts /e/ and /æ/, /u:/and /ʊ/as well as /ɔ:/ and /ɒ/ which are of CG type and also the minimal pairs with English vowel contrast /i:/and /ɪ/ which is of SC type. The PAM predicted the identification of TC type is the best, followed by CG type and the worst is SC type. It was hypothesized that the perceptual accuracy of English minimal pairs with vowel contrasts by the subjects in this study would follow the prediction by the PAM.

Methodology

Materials

An otoscope, tympanometer and a pure-tone audiometer were used for hearing screening. Four English front vowels, /i:/,/ɪ/,/e/ and /æ/ as well as five back vowels /u:/, /ʊ/, /ɔ:/, /ɒ/ and /ɑ:/ would be included as stimuli in the study. Each front vowel would be paired with the other three front vowels to form six minimal pairs. Similarly, each back vowel would be paired with the other four vowels to form ten minimal pairs. No mid vowels would be included in this study so as to make the categorization of front vowels and back vowels clear. As a result, sixteen minimal pairs with vowel contrasts would be introduced (Appendix A). The English stimuli used in both tasks could be found in the primary English textbooks in Hong Kong. All of them were consonant-vowel-consonant words, and were spoken in the accent of Received Pronunciation (RP). As pointed out by Chan (2012), the use of stimuli spoken in RP English may be regarded as inappropriate to the Hong Kong context due to the fact that this accent is not widely spoken in Hong Kong regardless of the fact that

Hong Kong was a colony of the United Kingdom where RP was a standard accent before 1997. However, there is no norm of Hong Kong English at the moment, and RP English is the most widely accepted accent for English teaching and learning (Chan, 2012) so it was selected in the perception tasks in the study. All the stimuli were recorded individually in a soundproof room using an industry-standard highly versatile cardioid dynamic vocal microphone SHURE SN58 and power amplifier M-AUDIO. All interstimulus intervals between stimuli were edited to 0.4 second using the computer programme "Audicity". A custom program for Apple Ipad Air was designed for presenting the stimuli to the participants via a Samsung EO-HS3303WEGWW earphone.

Participants

Thirty-two Hong Kong Cantonese EFL secondary school students were recruited for the study by convenient sampling, including 16 male and 16 female aged from 12 to 18. A hearing screening was performed. Otoscopic evaluation was conducted (Worrmald & Browning, 1996). Subjects had to possess type A tympanograms in both ear (Wong, Au & Wan, 2008) and a pure-tone air-conduction audiometry with threshold lower than 25dB HL at 500Hz, 1000Hz, 2000Hz, 4000Hz and 8000Hz in both ears ensure no hearing impairment of the participants (American National Standards Institute, 1978; American Speech-Language-Hearing Association, 1985; Martin & Clark, 2010). One male participant did not pass the hearing screening and thus was excluded from the research. A pre-test

language background survey was conducted (Appendix B). All the participants claimed to use Hong Kong Cantonese as their dominant language and English as a foreign language and have no cognitive or language problems. They were form one to form six students studying at eight local secondary schools, including 11 junior forms students (form one to three) and 20 senior forms students (form four to six). Twenty-seven participants studied in schools using English as the medium of instruction (EMI), and four participants studied in school using Chinese as the medium of instruction (CMI). All of them started to learn English formally at the age of six or younger. Twenty-three claimed to have received phonetic training in their English lessons by native English teachers, and the accent they learned was RP. Eight had not received any phonetics training before. In view of the number of years of English education received and the infrequent usage of English in daily situations, the whole group of participants could be regarded as low-to-intermediate level in English proficiency.

Procedures

The participants took part in a pre-task practice and two English vowel perception tasks. The recordings were presented using a custom program to the participants individually at a comfortable volume over earphones connected to an Apple iPad Air in a quiet room during the experiment. Responses were collected by tapping the iPad Air screen. The participants were asked to complete all the tasks in the same session. An assistant administered all the experiments and gave a short pre-test briefing about the experiment procedures to each

participant in Cantonese. This ensured that they understood the instructions clearly, thus enhanced the reliability and validity of the results.

Pre-task Practice. The written form and the corresponding pictures of the 32 stimuli of English words which would be presented in the following tasks were listed in iPad over two pages. Pronunciation of each stimulus was elicited by tapping the corresponding picture. After familiarizing with the stimuli, the 32 stimuli were played one by one, with random distracters of two other stimuli among the 32 stimuli. Each English word was spoken once without replay and at the same time three pictures with their corresponding name on the bottom would be displayed, and the participants were asked to tap on the picture or the written form of the stimuli they heard. There were 32 trials in total. Each stimulus was targeted once and used as the distracters of the other stimuli for twice. Having the pre-task practice was necessary to ensure that the participants were familiar with the test stimuli, as they were having low-to-intermediate but not advanced English proficiency.

Task 1 – English Vowel Discrimination. An ABX task was involved. In each trial, three English words of two stimuli were played once without replay. One of the two stimuli in a minimal pair (stimuli were only different in their vowels) was heard once and the other one in the same pair was heard twice. The order of them was randomized by the computer program so that the odd stimulus would be placed in different positions. The participants were then asked to choose the odd stimulus by tapping the number 1, 2 or 3 on the iPad

screen. All the stimuli appeared once so that there were 32 randomized trials.

Task 2 – English Vowel Identification. This was a spoken word-picture/written word (and picture) matching task, meant to examine the perceptual identification of different English vowels. Two pictures of minimal pairs with the corresponding words shown below were displayed on the iPad screen in each trial. The subjects were asked to tap on the picture of the word they heard without replay. Each stimulus had three trials so that there were 96 trials in total.

Results

Front Vowels

Figure 1 shows the results of discrimination of English front vowels. In this figure, the data are presented as the mean percentage accuracy of all participants. Discrimination between the minimal pairs with front vowel contrasts 1) /i/ and /ɪ/, 2) /i/ and /æ/, 3) /ɪ/ and /æ/ by the participants had 100% accuracy. The participants showed the poorest discrimination between the minimal pair with front vowel contrasts /ɪ/ and /e/ which had 87.1% accuracy. The results of all participants were above the significant level, tested using the ABX test ($N/2+\sqrt{N}$).

The results of a one-way repeated measures ANOVA indicated no significant main effect of accuracy of discrimination among the front vowel minimal pairs, Wilks' Lambda=.806, $F(3,28) = 2.24$, p=.106.

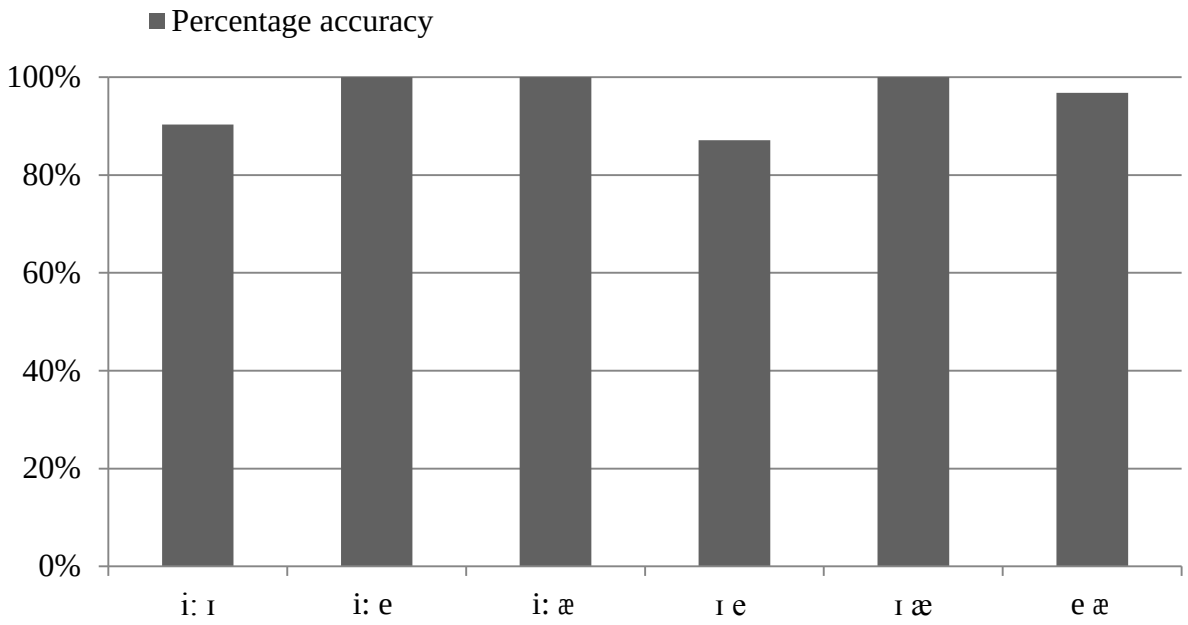

Figure 1. *Mean percentage accuracy of front vowel minimal pair discrimination in task 1*

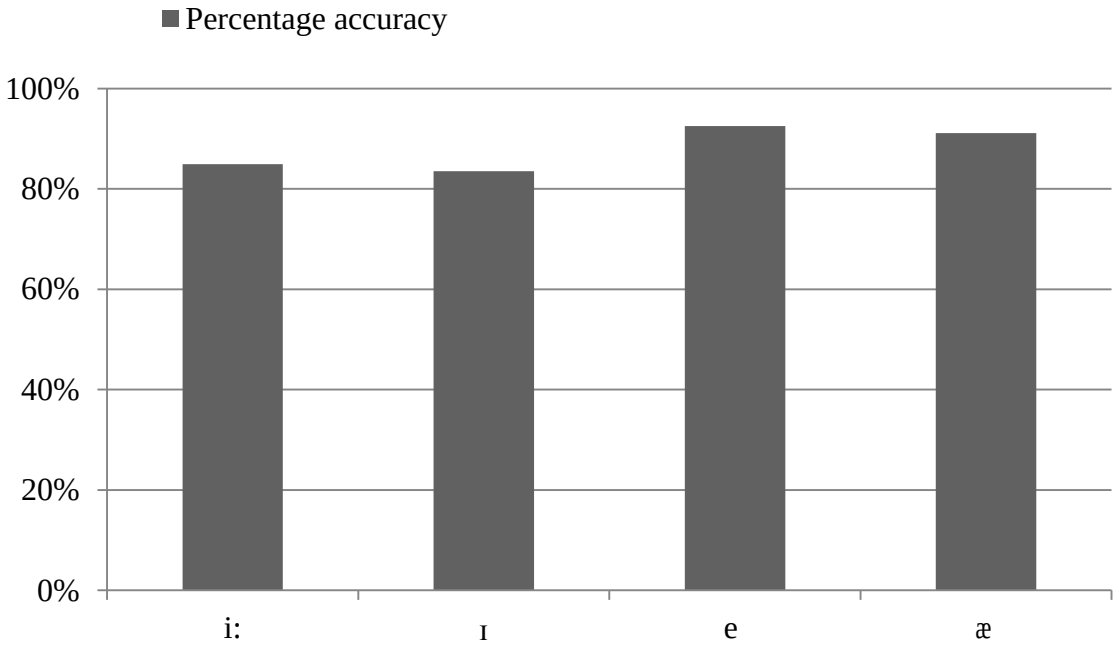

Figure 2. *Mean percentage accuracy of identification of front vowels in task 2*

Figure 2 shows the mean percentage accuracy of identification of front vowels in

task 2. Identification of the stimuli with front vowel /e/ achieved the highest accuracy of 92.5%, while that of the stimuli with /ɪ/ had the lowest accuracy of 83.5% which was higher than the chance level.

The results of a one-way repeated measures ANOVA indicated a significant main effects of accuracy of identification among the front vowels, Wilks'Lambda=.721, $F(3,28)=3.61$, $p=.026$. Post hoc comparisons using Bonferroni correction indicated that the pairwise difference between the stimuli with 1) /i:/ and /e/, 2) /ɪ/ and /e/ were significant($p<.05$), suggesting that identification front vowel /e/ is better than that of stimuli with front vowels /i:/ and /ɪ/.

Figure 3 shows the mean percentage accuracy of front vowel minimal pair identification in task 2. Identification of the minimal pairs with vowel contrast /ɪ/ and /æ/ attained the highest accuracy (98.4%), followed closely by the vowel contrasts /i:/ and /æ/ (97.8%) and /i:/ and /e/ (96.8%). Identification of the vowel contrast /i:/ and /ɪ/ was the worst (67.2%) among the all the front vowel minimal pairs which was higher than chance level.

A one-way repeated-measure ANOVA indicated significant main effects of accuracy of identification among the front vowel minimal pairs, Wilks' Lambda=.387, $F(5,26)=$, $p<.001$. Post hoc comparisons using Bonferroni correction indicated that identification between the minimal pair with front vowel contrast /i:/ and /ɪ/ was the worst among the front

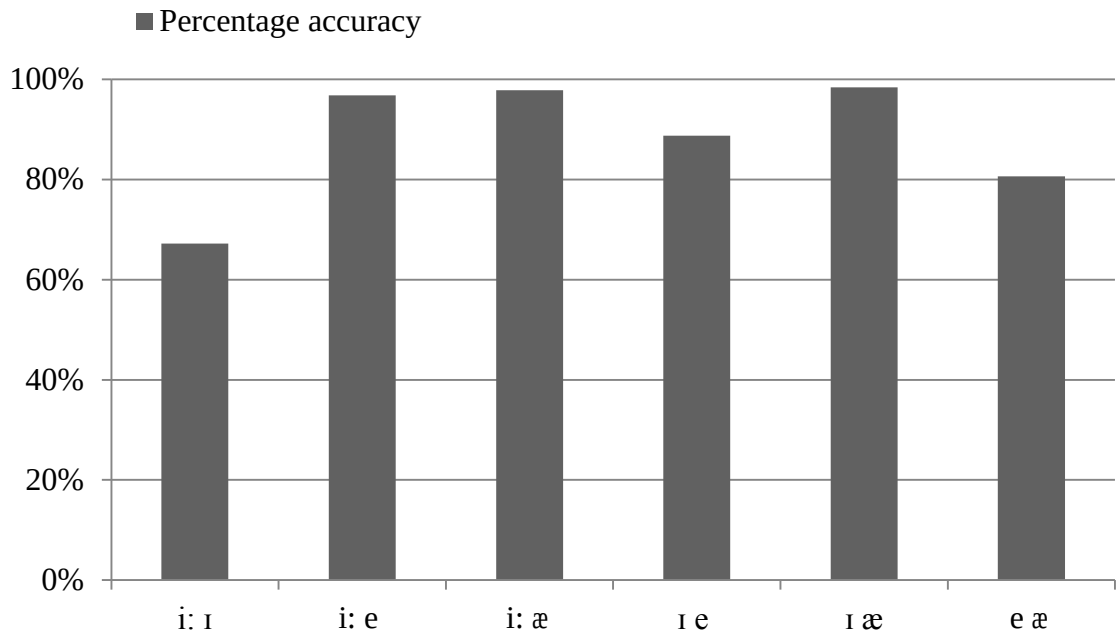

Figure 3. *Mean percentage accuracy of front vowel minimal pair identification in task 2*

vowel pairs and significantly poorer than that of the minimal pairs with front vowel contrasts 1) /i:/ and /e/, 2) /i:/ and /æ/, 3) /ɪ/ and /e/, 4.) /ɪ/ and /æ/ (p<0.05). Identification between the minimal pair with front vowel contrast /e/ and /æ/ was found to be the second worst and significantly lower than the minimal pairs with front vowel contrasts 1) /i:/ and /e/, 2) /i:/ and /æ/, 3) /ɪ/ and /æ/ (p<.05).

Table 1 shows the percentage accuracy of identification of the front vowel targets relative to other front vowel distracters. From table 1, the percentage accuracy of identification of front vowels /e/ with /ɪ/ as the distractor is 100% and is much higher than the other way round (77.4%). However, the difference was not statistical significant (p>.05), tested by paired-samples T-Test.

Table 1. *Mean percentage (%) accuracy of front vowel identification in task 2*

Distracters / Targets	Percentage accuracy of correct front vowel identification			
	i:	ɪ	e	æ
i:	--	58.1 (37.5)	96.8 (10.0)	100 (0)
ɪ	76.3 (31.3)	--	77.4 (26.4)	96.8 (12.3)
e	96.8 (10.0)	100 (0)	--	80.6 (29.5)
æ	95.7 (14.3)	100 (0)	80.6 (26.9)	--

Note. The figures in brackets represent the standard deviation of front vowel identification.

Back Vowels

Figure 4 shows the results of the discrimination Task of English back vowels. In this figure, the data are presented as the mean percentage accuracy of all the participants. Discrimination between the minimal pairs with back vowel contrasts 1) /u:/ and /ʊ/, 2) /u:/ and /ɔ:/, 3) /ʊ/ and /ɔ:/, 4.) /ʊ/ and /ɒ/, 5.) /ʊ/ and /ɑ:/, 6.) /ɔ:/ and /ɑ:/ by the participants had 100% accuracy. The participants showed the poorest discrimination between the minimal pair with back vowel contrasts /ɔ:/ and /ɒ/ which has 83.9% accuracy. The results of all the participants were above the chance level.

A one-way repeated measures ANOVA indicated no significant main effect of accuracy of discrimination among the back vowel minimal pairs, Wilks' Lambda=.817, $F(4,27) = 1.51$, p=.227.

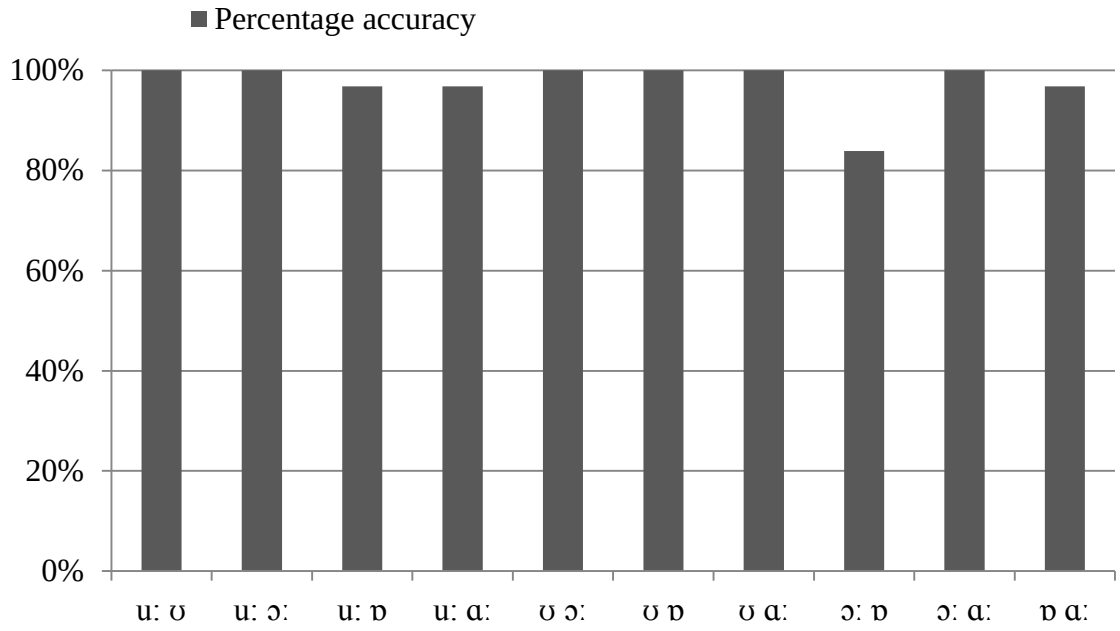

Figure 4. *Mean percentage accuracy of front vowel minimal pair discrimination in task 1*

Figure 5 shows the mean percentage accuracy of identification of back vowels in task 2. Identification of the stimuli with back vowel /ɑː/ achieved the highest accuracy of 95.7% while that of the stimuli with /ɒ/ had the lowest accuracy of 85.8% which was higher than chance level.

A one-way repeated measures ANOVA indicated significant main effects of accuracy of identification among the back vowels, Wilks' Lambda=.538, $F(4,27)=5.79$, $p=.002$. Post hoc comparisons using Bonferroni correction indicated that pairwise difference between 1) /u:/ and /ʊ/, 2) /u:/ and /ɑː/, 3) /ʊ/ and /ɒ/, 4) /ɒ/ and /ɑː/ were significant, $p<.05$, suggesting that the identification of English vowel /u:/ and /ɒ/ was worse than that of English vowel /ʊ/ and /ɑː/.

Figure 6 shows the mean percentage accuracy of back vowel minimal pair

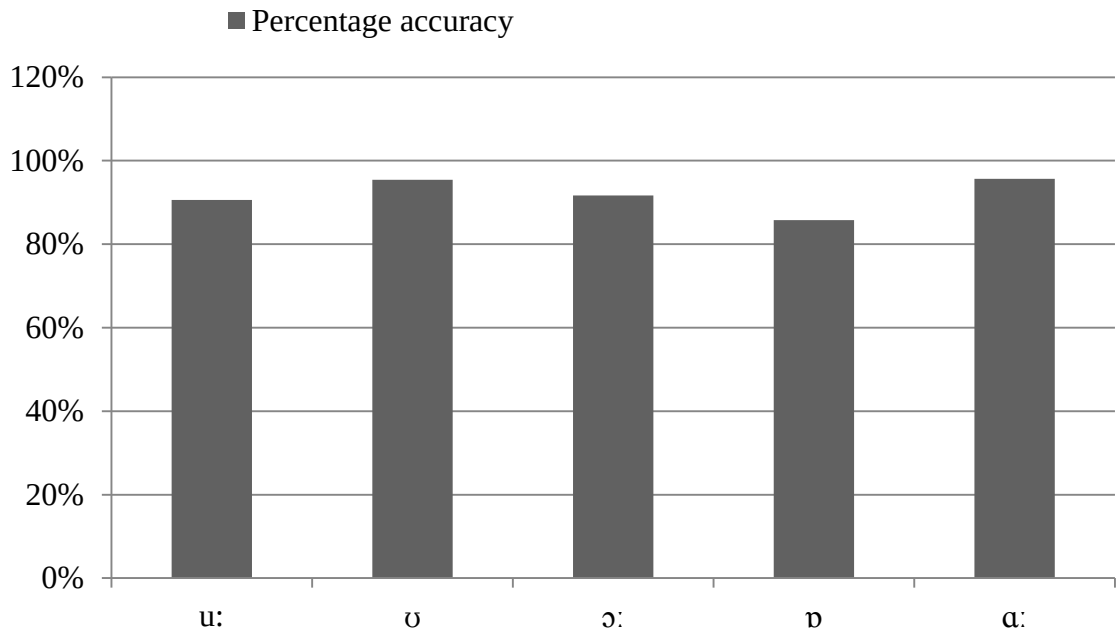

Figure 5. *Mean percentage accuracy of identification of back vowels in task 2*

identification in task 2. Identification between minimal pairs with vowel contrasts /ʊ/ and /ɒ/ as well as /ʊ/ and /ɑː/ attained the highest accuracy (99.5%), followed closely by those with vowel contrasts /uː/ and / ɑː/ (99.0%) as well as /ɔː/ and /ɑː/ (99.0%). Identification between the minimal pair with vowel contrast /uː/ and /ʊ/ (79%) was the worst among the back vowels.

A one-way repeated measures ANOVA indicated significant main effects of accuracy of identification among back vowel minimal pairs, Wilks' Lambda=.325, $F(9,22)=5.08$, $p=.001$. Post hoc comparisons using Bonferroni correction indicated that identification between the minimal pair with back vowel contrast /uː/ and /ʊ/ was the worst among the back vowel

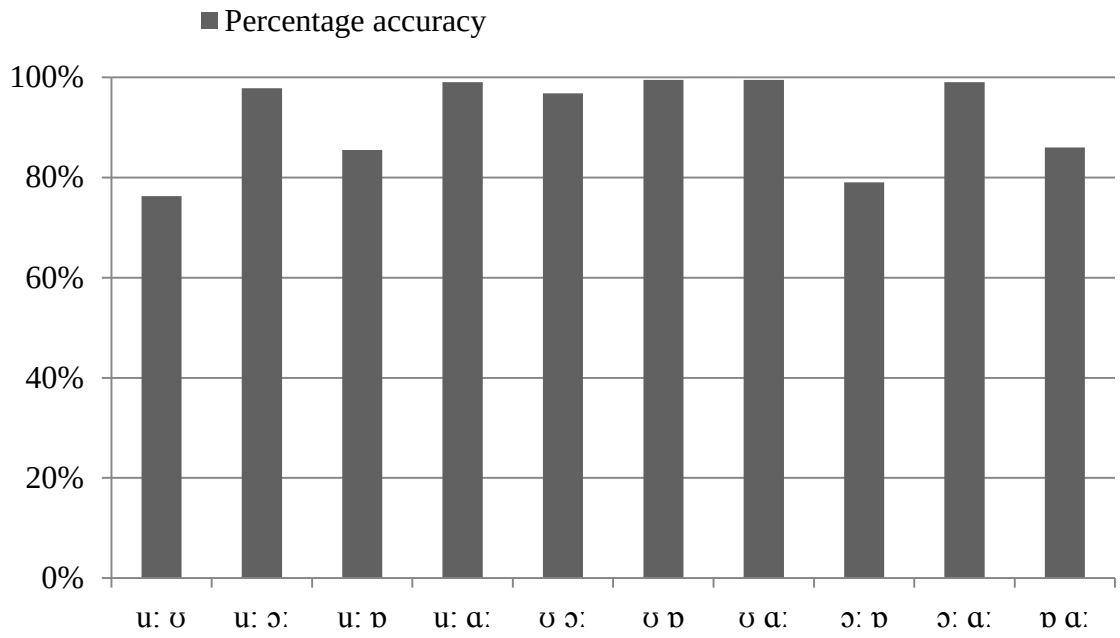

Figure 6. *Mean percentage accuracy of back vowel minimal pair identification in task 2*

pairs, followed by that between the minimal pair with back vowel contrast /ɔ:/ and /ɒ/. Both of them were significantly poorer than that of the minimal pairs with back vowel contrasts 1) /u:/ and /ɔ:/, 2) /u:/ and /ɑ:/, 3) /ʊ/ and /ɔ:/, 4) /ʊ/ and /ɒ/, 5) /ʊ/ and /ɑ:/, 6) /ɔ:/ and /ɑ:/ (p<0.05).

Table 2 shows the percentage accuracy of identification of the back vowel targets relative to other back vowel distracters.

Front Vowels vs. Back Vowels

A paired-samples T-Test indicated no significant main effect of accuracy for discrimination of minimal pairs with front vowel contrasts (M=2.96, SD=.083) and that with back vowel contrasts (M= 2.97, SD= .069); t (30) = 2.76, p=.675.

Table 2. *Mean percentage (%) accuracy of back vowel identification in task 2*

Distracters / Targets	Mean of correct back vowel identification				
	uː	ʊ	ɔː	ɒ	ɑː
uː	--	68.8 (29.7)	99.0 (6.0)	94.6 (17.4)	100 (0)
ʊ	83.9 (27.0)	--	99.0 (6.0)	99.0 (6.0)	100 (0)
ɔː	96.8 (10.0)	94.6 (15.1)	--	77.4 (30.3)	97.8 (12.0)
ɒ	76.3 (25.7)	100 (0)	80.6 (29.5)	--	86.0 (22.4)
ɑː	97.8 (8.3)	99.0 (6.0)	100 (0)	88.1 (20.3)	--

Note. The figures in brackets represent the standard deviation of back vowel identification.

Another paired-samples T-Test indicated significant main effects of accuracy for identification of stimuli with front vowels (M=5.30, SD=.596) and that with back vowels (M= 5.52, SD= .368); t (30) = 2.76, p=.010, suggesting that identification of back vowels was better than that of front vowels.

Discussion

Front Vowels

For front vowels, discrimination between the minimal pair with front vowel contrasts /ɪ/ and /e/ was the poorest. This result may be attributed to the fact that English vowels /ɪ/ and /e/ are the closest among English front vowels in terms of backness and height of the vowels (Chan & Li, 2000). However, there is no significant main effect of accuracy of

discrimination among the front vowel minimal pairs.

Besides, the results of task 2 show that identification of English front vowel /e/ attained the highest accuracy among the front vowels and is significantly higher than the English front vowels / i:/ and /ɪ/ (p<0.05). This can be explained by the fact that the English vowel /e/ has the most similar assimilated Cantonese vowel /e/ compared to other English front vowels (Chan & Li, 2000).

According to the PAM, perception of the minimal pair with front vowel contrast /i:/ and /ɪ/ ,which is of SC type, is expected to be poor while that of the minimal pair with front vowel contrast /e/ and /æ/, which is of CG type, is expected to be moderate to good. The other four minimal pairs with front vowel contrasts are of TC type, the perception is expected to be excellent and better than the two mentioned pairs.

From the result, identification of minimal pair with front vowel contrast /i:/ and /ɪ/ is found to be the worst among all the front vowel pairs and significantly poorer than the four minimal pairs with front vowel contrasts of TC type assimilation (p<0.05). Identification of minimal pair with vowel contrast /e/ and /æ/ is found to be moderate and significantly lower than the minimal pairs with front vowel contrasts /i:/ and /e/, /i:/ and /æ/ as well as /ɪ/ and /æ/ which are of TC type assimilation (p<0.05). The findings support the prediction of the PAM that perception of minimal pairs with TC type contrast is the best and that of minimal pairs with SC type contrast is the worst.

It is also found that the identification of the minimal pair with front vowel contrast /ɪ/ and /e/ is the worst among the minimal pairs with TC type assimilation and is significantly lower than the other three minimal pairs with TC type assimilation (p<0.05). For instance, the accuracy of identification of /e/ with /ɪ/ as the distracter was significantly higher than that of discriminating /ɪ/ with /e/ as the distracter. According to Chan and Li (2000), English vowel /e/ is highly similar to Cantonese vowel /e/ while English vowel /ɪ/ is between Cantonese vowels /e/ and /i/ in terms of height of pronunciation. It is possible that when the participants listened to the stimuli with English vowel /ɪ/, they would struggle of assimilating English vowel /ɪ/ into Cantonese vowel /e/ or /i/, resulting in poorer identification.

Back Vowels

For back vowels, the poorest discrimination between the minimal pair with back vowel contrasts /ɔː/ and /ɒ/ is found. This result may be attributed to the findings by Chan (2012) that English vowels /ɪ/ and /e/ are both assimilated to Cantonese vowel /ɔ/, leading to poor discrimination. However, there is no significant difference in participants' percentage accuracy of discrimination among the back vowel minimal pairs.

Besides, the results of task 2 show that identification of English back vowel /ɑː/ attained the highest accuracy among the back vowels and is significantly higher than English back vowels /uː/ and /ɒ/ (p<0.05). The can be explained by the fact that English vowel /ɑː/ has a more similar assimilated Cantonese vowel /aː/ compared to English front vowels /uː/

and /ɒ/ (Chan & Li, 2000).

According to the PAM, identification accuracy of the minimal pairs with back vowel contrast /u:/ and /ʊ/ as well as /ɔ:/ and /ɒ/ which are of CG type, are expected to be poor than the other eight back vowel pairs which have TC type, and their identification is expected to be excellent.

From the result, the identification of back vowels /u:/ and /ʊ/ as well as /ɔ:/ and /ɒ/ was the worst among the back vowel pairs and was significantly poorer than most of the pairs with TC type assimilation. As pointed out by Wang (2002), there is a loss of clear two-category distinction of English back vowels /u:/ and /ʊ/ for Mandarin speakers. This may also apply to Cantonese speakers and explain the relatively poorer result of discriminating this pair than other back vowels pairs. However, their identification was better than the front vowels /i:/ and /ɪ/ pair which has the SC type assimilation. The findings again support the prediction of the PAM that perception of minimal pairs with TC type contrast is the best, followed by CG type, and that of minimal pairs with SC type contrast is the worst.

The identification of the /u:/ and /ɒ/ as well as the /ɒ/ and /ɑ:/ back vowel pairs were not significantly better than the worst pair /u:/ and /ʊ/ ($p > .05$) although they are of TC assimilation type. For the /u:/ and /ɒ/ vowel pair, it is highly possible that the semantic-related stimuli “bomb” and “boom” could account for the worse result to other TC

pairs, as pointed out by some participants who found reported that they were still confused with the pronunciation and meaning of /boom/ and /bomb/ after taking the test.

From the overview of English vowels and Cantonese vowels (Chan & Li, 2000), both Cantonese vowels /ɔ/ and /a:/ are similar to /ɒ/ regarding the backness and height of the vowels, thus it is possible for listeners to assimilate English vowel /ɒ/ to Cantonese vowel /a:/ as well as /ɔ/. In this case, English vowel pair /ɒ/ and /ɑː/ will be of CG type assimilation. As a result, the identification performance of this pair lies between the TC assimilation and CS assimilation, thus there was no significant difference of this pair to any other pairs which are of TC and CS type. If this is the case, English vowel /ɒ/ is sometimes assimilated to Cantonese vowel /a:/; then English vowel pair /ɔ/ and /ɒ/ can also be TC type assimilation which is expected to have excellent perceptual result. When comparing the identification result of the minimal pair with vowel contrasts /ɔ/ and /ɒ/ with that of /ɒ/ and /ɑː/, the poorer result of identifying English vowel pair /ɔ/ and /ɒ/ relative to /ɒ/ and /ɑː/ implied that English vowel /ɒ/ is more often be assimilated to Cantonese vowel /ɔ/.

Limitations

This study had a few limitations. First, as the subjects of this study were recruited only from local secondary schools, the result may not be generalized to other Cantonese EFL learners with different English proficiency. Also, the effect of linguistic background such as the medium of school instruction was not investigated due to the large difference between the

sample sizes of the two groups. Besides, the minimal pair "boom" and "bomb" are semantically related, affecting the validity of the result. Moreover, only vowels but no consonants were included in the study. Last but not least, the PAM could not provide explanation for the significant difference between the identification of stimuli with front vowels and that of back vowels.

Conclusion

This study investigated Hong Kong Cantonese EFL listeners' perceived relations between English and Cantonese speech sounds. The applicability of the PAM to the foreign language phonology acquisition was examined. It is found that the identification of minimal pairs with vowel contrasts that have TC type assimilation has the highest accuracy, followed by CG type, and the worst is minimal pairs with SC assimilation type vowel contrast. This provides support to accept the hypothesis that the perceptual accuracy of English minimal pairs with vowel contrasts by the subjects in this study would follow the prediction by the PAM. However, given that the subjects are all local secondary school student in Hong Kong, the results cannot be generalized to include all learners in Hong Kong at different proficiency levels.

Further research is needed to include English learners from different language proficiency such as the English learning beginners to have a more comprehensive view of perception of English vowels of Hong Kong Cantonese EFL learners. Inclusion of

comparable number of participants with different language backgrounds should also be done to investigate the effect of language background. Perception of English consonants is also needed to be investigated.

Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Lena Wong for giving me generous guidance and supervision from the beginning till the end. I would also like to thank Dr. Felix Chen for the technical assistance on designing the custom IPad program for the experiment.

Other than academic support, I would like to express my sincere gratitude to the my friends, Tszyan, b, Ngating, Florence, Chingyee, Claudia and Jenny as well as the teaching staffs in my secondary school, in particular, the Principal Mr. Tam, teachers Ms. Siu, Mr. Tong, Mr. Fong, Ms. Ho and Mr. Tsui for their valuable help in recruiting the subjects.

Last but not least, my family's emotional support was of paramount importance to me. Without their unlimited tolerance and care for me, the study would not be accomplished.

References

American National standard Standards Institute. (1978). M*ethods for manual pure-tone threshold audiometry.* ANSI S3.21-1978, R-1986. New York: ANSI.

American Speech-Language-Hearing Association. (1985). *Guidelines for identification*

audiometry. ASHA, 27, 49-53.

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., McRobert, 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. doi:10.1121/1.1332378

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

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.

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.

Guion, S., Flege, J. E., Akahamada, 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*, 2711-2724.

Kingston, J. (2003). Learning foreign vowels. *Language and Speech, 46*, 295-349.

Martin, F. N., & Clark, J. G. (2010). *Introduction to Audiology.* (10th ed.). Boston: Allyn and Bacon.

Wang, X. (2000). *Training Mandarin and Cantonese speakers to discriminate English vowel contrasts: Long-term retention and effect on production.* (Doctoral dissertation). Retrieved from http://zimmer.csufresno.edu/~xinw/contents.pdf

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. Doi: 10.1097/AUD.0b013e318164aa61

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

Appendix A

Stimuli used in task 1 and task 2

Vowel contrasts		Minimal pairs	
i:	ɪ	Sheep	Ship
i:	e	Read	Red
i:	æ	Seed	Sad
ɪ	e	Pin	Pen
ɪ	æ	Hit	Hat
e	æ	Bed	Bad
ɔ:	a:	Four	Far
ɐ	a:	Hot	Heart
u:	ʊ	Pool	Pull
u:	ɔ:	Cool	Call
u:	ɐ	Boom	Bomb
u:	a:	Juice	Jars
ʊ	ɔ:	Full	Fall
ʊ	ɐ	Good	God
ʊ	a:	Book	Bark
ɔ:	ɐ	Forks	Fox

Appendix B

Demographic Questionaire

姓名：_____

班級：_____

出生日期：_____

出生地點：_____

母語：_____

與家人/照顧者溝通的主要語言（可多於一種）：_____

除香港以外曾否在外地定居（如有，請註明）：_____

除學校以外有否接觸/使用英語的機會（如有，請註明）：_____
_____

曾否接受正式英語拼音訓練（如有，請註明）：_____
_____

有否已確診的言語障礙或言語發展遲緩（如有，請註明）：_____
_____