

# The Effects of High and Low Variability Phonetic Training on the Perception and Production of English Vowels /e/-/æ/ by Cantonese ESL Learners with High and Low L2 Proficiency Levels

Janice Wing Sze Wong

Hong Kong Baptist University

janicewong@hkbu.edu.hk

## Abstract

This study investigated the effects of two perceptually-based training paradigms in both the perception and production of English /e/ and /æ/ by Cantonese ESL learners with high and low listening and oral proficiency levels. Sixty-four subjects participated in the study, in which 22 (9 with high proficiency, H-HV; 13 with low proficiency, L-HV) were trained under High Variability Phonetic Training (HVPT) approach, 19 (8 with high proficiency, H-LV; 11 with low proficiency, L-LV) were trained under Low Variability Phonetic Training (LVPT) approach whereas 23 (10 with high proficiency, H-CO; 13 with low proficiency, L-CO) were the control subjects. Both training approaches were effective in improving the subjects' perception of the two vowels, with HVPT groups showing more robust improvement than LVPT groups. Perceptual learning could also be generalized to new words and new speakers and be transferred to the production domain, with HVPT groups outperforming LVPT groups. However, subjects with different proficiency levels learned to similar degrees in all tests. The results demonstrated that both approaches offered a type of learning that allows attention to focus on phonetic information, which is different from what is learned in an L2 classroom; whereas stimulus variability also plays a role in the learning.

**Index Terms:** High Variability Phonetic Training, L2 acquisition, perception and production of non-native contrast, L2 proficiency

## 1. Introduction

Previous studies have shown that phonetic training could be useful for modifying the perception and production of some non-native contrasts [1]-[3]. Among the many training methods, High Variability Phonetic Training (HVPT), which involves the use of natural training stimuli with various phonetic contexts produced by multiple speakers with immediate feedback, has received particular attention and studies have reported its effectiveness and suggested that stimulus variability is the core element in making the approach useful [4]-[16]. However, most of them had not compared directly with subjects trained with low variability stimuli to provide more solid evidence to the above claim [c.f. 17]. Meanwhile, the above training studies recruited subjects at similar proficiency levels. It would be insightful if subjects with different proficiency levels could be compared since it could be logical that high proficiency learners would not benefit too much from the training since they already had better perception and production abilities, or even possibly more L2 experience than the low proficiency ones. This study aimed to discover whether phonetic training could benefit subjects with different proficiency levels and to learn about how the training approaches affected the learning. L2 teachers and learners can also benefit from the present research. The

present study has chosen to investigate the perception and production of a commonly-confused vowel pair /e/-/æ/ [18]-[20], among Hong Kong Cantonese learners of English with high and low listening and oral proficiency levels.

## 2. Methodology

### 2.1. Participants

A total of 64 native Hong Kong Cantonese speakers (average age = 16.97) with English as their L2 participated in the study. They were put into different groups according to the training type and their L2 proficiency levels determined by the average of listening and oral exam of a locally and internationally recognized public exam in Hong Kong [20]. 22 (9 with high proficiency, H-HV; 13 with low proficiency, L-HV) were trained under High Variability Phonetic Training (HVPT) approach, 19 (8 with high proficiency, H-LV; 11 with low proficiency, L-LV) were trained under Low Variability Phonetic Training (LVPT) approach whereas 23 (10 with high proficiency, H-CO; 13 with low proficiency, L-CO). They all started learning English as an L2 for 13.94 years (SD = .91). They had not resided in any English-speaking countries and none reported hearing or speaking deficits.

Seven other native speakers of English (4 females 3 males) were also invited to produce perceptual stimuli for the tests and training. Their ages ranged from 20 to 60, all speaking General American accent.

### 2.2. Design

All groups of subjects underwent the three phrases:

PHASE 1. Pretest Phase, which included one production pretest and one perception pretest;

PHASE 2. Treatment Phase:

- High Variability Phonetic Training, with a total of 10 training sessions (only the H-HV and L-HV groups participated) in which they attended two sessions a day, each lasted for 10 minutes;
- Low Variability Phonetic Training which was different from HVPT only in the use of stimuli (only the H-LV and L-LV groups participated)

PHASE 3. Posttest Phase, with one production posttest, one perception post-test, one production Test of Contextualization (TC) and two perception Tests of Generalization (TG1, TG2).

### 2.3. Setting and Apparatus

All the subjects completed all perceptually-related tests and/or training sessions by a computer program designed by the researcher in a language laboratory. The subjects listened to perceptual stimuli presented in the program and data were saved into a Microsoft Access database. They also recorded their production test tokens using Adobe Audition 1.5.

## 2.4. Stimulus Materials

All the perceptual pre/posttest and training stimuli were made by six of the seven native English speakers. Each speaker produced 10 pairs of /e/-/æ/ real word pairs, contributing a total of 60 stimuli with a wide variety of phonetic environments (with different CVC and syllable structures). These were for the HVPT sessions. For the LVPT, one of the six speakers recorded all 60 word pairs. All the productions made by this speaker were also utilized in all the perception tests. Also, among the six native speakers, i.e., a familiar speaker to the subjects, also recorded a word list for TG2 (new words by a familiar speaker) which included 30 new /e/-/æ/ minimal pairs. The last speaker who had not recorded anything for the training or the tests, known as a new speaker, recorded another new list with 30 /e/-/æ/ minimal pairs for TG1 (new words by a new speaker). Each speaker read the tokens at least three times to avoid intra-speaker variability.

## 2.5. Procedure

### 2.5.1. Pretest Phase

All subjects participated in this phase in which there were two tests: a production pretest and a perception pretest. The production pretest was completed before the perception pretest to avoid any cueing effects.

- *Production pretest:* The subjects were given a word list of 60 words (30 /e/ and 30 /æ/) and 10 distractors which might appear in the perception tests or the training. They did five practice trials first and were asked to produce the tokens with natural loudness and speaking rate. No audio prompts or instructions were given during the recording. They could also pause and resume the recording based on their own pace. This lasted for less than 10 minutes.

- *Perception pretest:* The subjects completed the test within 30 minutes. There were 70 questions in total (60 words with either /e/ or /æ/ plus 10 distractors). Five practice trials were completed before the test. Before they confirmed their answers, they listened to the stimuli and chose the answer from three choices with conventional English orthography, or a blank for free answer for them to type their own word. The frequency of occurrence of the correct answer appeared in the four serial positions, i.e. word 1, word 2, word 3, free answer, were equal, allowing the chance level to be correctly and fairly inferred at 25%. This design was aimed to reveal a more genuine and reliable performance of the subjects by avoiding the 50% probability of correct answer obtained in typical two-alternative identification test with only two choices.

### 2.5.2. Training Phase

Only the four experimental groups participated in this phase. They took two training sessions per day for 5 days. A 30-minute break with refreshments was given between sessions.

- *The HVPT:* A total of 60 stimuli produced by six different native English speakers were presented to the subjects. All the tokens were randomized in terms of speakers and order. The subjects were trained on a two-alternative forced choice paradigm so as to raise the intensiveness of the training effect and reduce the cognitive load when choosing among more options. They can pay more attention to and focus on identifying only the two vowels perceptually without interference from other sounds. The stimuli were one of the counterparts in a minimal pair contrasting the two vowels (e.g.

either “bed” or “bad”). During training, immediate feedback was given; at the end of each session, their total scores were also shown for track-keeping of the training progress.

- *The LVPT:* The training procedures were exactly the same as in the HVPT. The only difference was in the training stimuli: the same 60 words were produced by only one female native speaker of English.

### 2.5.3. Posttest Phase

All subjects participated in this phase which involved two production posttests (posttest and TC) which were completed before the three perception posttests (posttest, TG1 and TG2).

- *Production posttest:* same as the production pretest.

- *Test of Contextualization (TC):* All the subjects were given a 250-word passage which included 50 content words with /e/ and /æ/. They were asked to produce the whole passage naturally, at their own pace and loudness.

- *Perception posttest:* same as the perception pretest

- *Test of Generalization 1 (TG1):* The subjects heard 30 new words spoken by a new speaker. The procedures were similar to those administered in the perception pretest.

- *Test of Generalization 2 (TG2):* The subjects listened to 30 new words spoken by a familiar speaker. Procedures were the same as those in TG1.

## 2.6. Evaluation of Production Data

All production test tokens were transcribed twice by a phonetically-trained researcher with Cantonese as L1 and English as L2. The correlation coefficient of the two transcriptions was computed and the intra-rater judgment was obtained at  $r = .982$  ( $p < .001$ ). Another phonetically-trained doctoral student studying Applied English Linguistics with Cantonese as L1 and English as L2 also transcribed the data phonetically. The reliability check was done without referring to any completed transcriptions and it was 94.74%. The intra-rater reliability was obtained as  $r = .907$  ( $p < .001$ ). A follow-up acoustic analysis on  $F1$ ,  $F2$ ,  $F3$  values and the vowel durations also reinforced the transcriptions.

## 3. Results

### 3.1. Perceptual Performance

#### 3.1.1. Effects of training: Pretest vs. Posttest

This boxplot displays the results of subjects (according to the training type only) in the pretest and posttest:

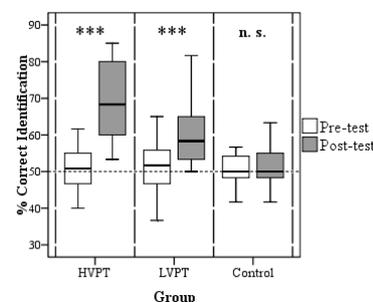


Figure 1. Mean percentages of correct identification of the three groups in the pretest (white boxes) and the posttest (dark boxes) [\*\*\* =  $p < .001$ ; n.s. =  $p > .05$ ]. The dashed line indicates the chance level performance.

A four-way repeated measures ANOVA was computed using Group (HVPT, LVPT, Control), Test (pretest, posttest), Vowel (/e/, /æ/) and Proficiency (high, low) as factors. It showed highly significant main effects of Group [ $F(2,58) = 11.92, p < .001$ ], Test [ $F(1,58) = 60.23, p < .001$ ] and Vowel [ $F(1,58) = 19.55, p < .001$ ]; yet, Proficiency was not a significant factor ( $p = .595$ ). The interactions Group  $\times$  Test [ $F(2,58) = 25.11, p < .001$ ], Group  $\times$  Vowel [ $F(2,58) = 8.47, p = .001$ ], Test  $\times$  Vowel [ $F(1,58) = 7.44, p = .008$ ] and Test  $\times$  Vowel  $\times$  Proficiency [ $F(1,58) = 6.03, p = .017$ ] were all significant. Planned comparisons with Bonferroni correction on Group  $\times$  Test interaction showed, besides both HVPT and LVPT groups improved significantly from pretest to posttest (both at  $p < 0.001$ ), a significant difference between groups in the posttest, but not in the pretest ( $p = 1.00$ ), was also found. In the posttest, HVPT outperformed LVPT by 9.91% ( $p = 0.02$ ) and the control group by 19.18% ( $p < 0.001$ ) whereas LVPT also outperformed the control group by 9.27% ( $p = 0.04$ ).

### 3.1.2. Generalizability of training

The following two boxplots show the results in TG1 (left) and TG2 (right):

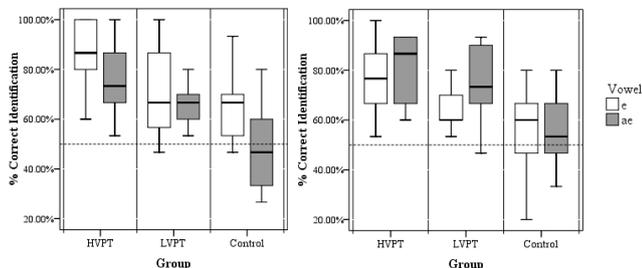


Figure 2. Mean percentages of correct identification of the three groups in TG1 (left) and TG2 (right) across groups. The dashed line indicates the chance level performance.

Two separate three-factors ANOVA with Group (HVPT, LVPT, control) and Proficiency (High and Low) as between factors and Vowel (/e/, /æ/) as a within factor were conducted for TG1 and TG2. For TG1, a significant effect of group [ $F(2,58) = 23.87, p < .001$ ] was obtained with HVPT  $>$  LVPT (mean difference = 12.99%;  $p = .006$ ); HVPT  $>$  Control (mean difference = 26.13%;  $p < .001$ ); LVPT  $>$  Control (mean difference = 12.14%;  $p = .005$ ). A robust effect of vowel [ $F(1,58) = 31.16, p < .001$ ] was also found, with /e/ being substantially more accurately identified than /æ/. The Vowel  $\times$  Proficiency interaction was also significant [ $F(1,58) = 8.08, p = .006$ ]. Post-hoc pairwise comparison (Bonferroni) indicated that both the high [ $F(1,61) = 32.95, p < .001$ ] and low [ $F(1,58) = 7.20, p = .010$ ] proficiency groups identified the vowel /e/ more accurately than /æ/ in the test. Nonetheless, the effect of proficiency and the interactions of both Group  $\times$  Vowel and Group  $\times$  Proficiency were not significant.

For TG2, the main effect of group was significant [ $F(2,61) = 27.10, p < .001$ ]. Yet, the effect of Proficiency and the interactions of Vowel  $\times$  Group, Group  $\times$  Proficiency, and Vowel  $\times$  Group  $\times$  Proficiency were not significant. The HVPT group again outperformed the LVPT group with significance (mean difference = 8.96%;  $p = .033$ ); it also had a higher mean scores than the control group (mean difference = 23.14%;  $p < .001$ ), and so did the LVPT over the control group (mean difference = 14.17%;  $p < .001$ ). Meanwhile, Vowel was also a

slightly significant effect [ $F(1,61) = 4.89, p = .031$ ], due to a more general accurate identification of the vowel /e/ than /æ/. Vowel  $\times$  Proficiency interaction was significant [ $F(1,61) = 9.86, p = .003$ ] and simple effects test showed that the identification of the vowel /e/ was more accurate than /æ/ in the low proficiency group and this interaction had significance. In addition, this group was also performing significantly better than the high proficiency one ( $p = .027$ ) in identifying vowel /e/.

## 3.2. Transfer of perceptual learning to production

### 3.2.1. Effects of training: Pretest vs. Posttest

This boxplot displays the results of production pretest versus posttest across training types:

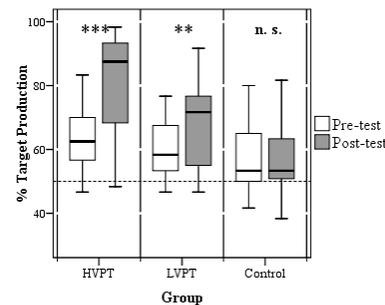


Figure 3. Percentage of target production of the three groups between the pretest (white boxes) and the posttest (dark boxes) [\*\*\* =  $p < .001$ ; \*\* =  $p < .01$ ; n.s. =  $p > .05$ ]. The dashed line indicates the chance level performance.

A four-way repeated measures ANOVA was computed using Group (HVPT, LVPT, Control), Test (pretest, posttest), Vowel (/e/, /æ/) and Proficiency (high, low) as factors. Significant main effects of Group [ $F(2, 58) = 8.63, p = .001$ ], Test [ $F(1, 58) = 93.77, p < .001$ ] and Vowel [ $F(1, 58) = 146.30, p < .001$ ] were found. Proficiency was not a significant factor though ( $p = .05$ ). The interactions Group  $\times$  Test [ $F(2,58) = 32.05, p < .001$ ], Proficiency  $\times$  Test [ $F(2,58) = 10.34, p = .002$ ], Group  $\times$  Vowel [ $F(2,58) = 10.08, p < .001$ ], Test  $\times$  Vowel [ $F(1,58) = 6.93, p = .011$ ], Test  $\times$  Group  $\times$  Proficiency [ $F(2,58) = 3.52, p = .036$ ] and Test  $\times$  Vowel  $\times$  Proficiency [ $F(1,58) = 15.18, p < .001$ ] were all significant. Planned comparisons with Bonferroni correction on Test  $\times$  Group  $\times$  Proficiency interaction showed that both H-HV and L-HV groups improved their production of the two vowels from pretest to posttest for 21.48% and 15.90% respectively ( $p < .001$ ), whereas only H-LV group had a robust 15.00% of improved by 15.00% ( $p < .001$ ) but not for L-LV group ( $p = .308$ ). However, the difference between the accurate production of the two vowels in the posttest for H-HV and L-HV groups was not significant ( $p = .289$ ); only the difference between H-LV and L-LV was significant ( $p = .037$ ). Meanwhile, concerning the difference between subjects with different proficiency levels trained under different paradigms, the results showed that H-HV only performed significantly better than H-CO (for 22.72%,  $p = .001$ ), but not the H-LV group in the posttest ( $p = .078$ ). For the performance of the low proficiency subjects in the posttest, L-HV outperformed both L-LV (by 15.06%,  $p = .02$ ) and L-CO group (by 23.08%,  $p < .001$ ).

### 3.2.2. Generalizability of training

This figure shows the results of TC across groups:

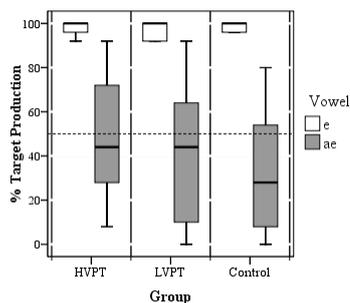


Figure 4. Percentage of target production of the three groups in TC. The dashed line indicates the chance level performance.

A three-factor ANOVA with Group (HVPT, LVPT, Control) and Proficiency (High and Low) as between-subjects factors and Vowel (/e/, /æ/) as a within factor demonstrated a significant effect of Proficiency [ $F(1,58) = 6.24, p = .015$ ], as the high proficiency group produced the target vowels more accurately than the low proficiency group. There was also a significant effect of Vowel [ $F(1,58) = 181.86, p < .001$ ], attributed to the huge difference in the production performance of the vowel /e/, which was over 95% and the vowel /æ/, which was only around 40% in average. Yet, the effect of Group was not robust ( $p = .062$ ). Among all interactions, only a very marginal significant interaction Vowel  $\times$  Proficiency [ $F(1,58) = 4.08, p = .048$ ] was found. A test of simple effects revealed that the differences between vowels were highly significant in both proficiency groups: [ $F(1,58) = 56.80, p < .001$ ] for high and [ $F(1,58) = 142.64, p < .001$ ] for low, due to the fact that the vowel /e/ was always more accurately produced than the vowel /æ/. While considering the differences between vowels, it was found that significant difference between proficiency groups was only observed in the vowel /æ/ [ $F(1,58) = 5.33, p = .009$ ] but not in the vowel /e/ ( $p = .924$ ), meaning that the two proficiency groups performed particularly similar in the production of /e/.

## 4. Discussion

The results of the present study have provided further support to the efficacy of the phonetic training approaches, both HVPT and LVPT in modifying the perception and production of non-native contrast /e/ and /æ/ by Cantonese ESL learners with high and low proficiency levels. HVPT was also a more effective approach than LVPT in terms of the improvement in the perception and production of the two vowels among the subjects. It infers that perceptual stimuli produced by multiple talkers were more useful in modifying the perceptual acquisition and (thus) production of the subjects, which is in line with earlier studies showing the effective and successful adoption of the HVPT in the modification of different sound segments. The general success of both HVPT and LVPT may be due to the consistent use of the same task which led to a consistent mapping of the phonological categorization through referencing between the training stimuli and the responses. The adoption of feedback might also have played a role.

Although previous studies have reported the importance of adopting multi-speaker highly-variable stimuli so as to successfully training the learning of a non-native contrast,

many of them could not support the claim since they had not compared their results with a group receiving stimuli with low variability. This study compared HVPT and LVPT directly and provided more solid evidence on the plausibility that it was the highly variable stimuli that led to the more successful learning of the non-native contrasts. Stimuli with higher variability could promote selective attention of the subjects by avoiding idiosyncrasies which may be brought by stimuli produced just by one speaker as well as providing a wider range of stimuli for the subjects to stretch the exemplar perceptual space wider along dimensions where the two vowels differ and be shrunk along dimensions that do not show distinctions for the two vowels. Also, HVPT provided a better simulation of real-life experience than LVPT as more voices could be heard during training.

Another preliminary finding in this experiment is that now high and low proficiency learners who learnt and were judged based on their general perception and production abilities generally learnt to a similar extent after training. It suggests that only through this type of laboratory and intensive training can L2 learners be able to learn some difficult non-native contrasts which cannot be learnt through an L2 classroom. It further implies that this type of phonetic training in the form of an identification task could direct the attention of the learners to the phonetic differences which improve L2 perception (and/or production). It infers that this type of training is totally different from common language learning in an L2 classroom, and that general perception and production abilities and experience are another level of underlying processes and representations that do not equal or cannot determine the perception ability at the auditory level.

## 5. Conclusion

These above findings showed that exposing learners to highly-variable natural stimuli can successfully train the perception and production of a non-native phonetic contrast among Cantonese speakers of English. Positive generalization effects to the perceptual domain suggest the efficacy of the HVPT is externally valid. Also, a solid gain from the training can be seen on both low and high proficiency Cantonese ESL learners. It suggests that the HVPT is not only effective for highly advanced learners as previously reported, but learners with lower listening and speaking proficiency can improve as well as the higher proficiency ones. This study fills the research gap by enriching relevant perceptual training studies concerning Cantonese speakers of English and offering more solid conclusion to the claim that stimulus variability is at work under laboratory-based perceptual training by comparing and contrasting directly the HVPT and the LVPT paradigms. Meanwhile, both high and low proficiency learners could gain similarly from the training may suggest that this type of intensive training may have provided learners with different aptitudes with a subtle type of training that can direct their attention to the difficult contrast and their acoustic properties in particular. This is certainly distinct from the learning of one way of understanding speech (perception) and presenting ideas (production) in ordinary L2 classrooms. Future research can study the effects of different orientations of attention and the efficacy of different paradigms.

## 6. References

- [1] Best, C. T., "A direct realist perspective on cross-language speech perception," in *Speech Perception and Linguistics Experience: Issues in Cross-Language Research*, W. Strange, Ed., Timonium, MD: York Press, 1995, pp. 171-206.
- [2] Flege J. E., "Factors affecting degree of perceived foreign accent in English sentences." *J. Acoust. Soc. Am.*, vol. 84, pp. 70-79, 1988.
- [3] Leather, J. Interrelation of perceptual and productive learning in the initial acquisition of second-language tone. In A. James & J. Leather (Eds.), *Second-language speech: Structure and Process*. Berlin: De Gruyter, 1997, pp. 75-101.
- [4] Bradlow, A., Pisoni, D., Akahane-Yamada, R., and Tohkura, Y., "Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production," *J. Acoust. Soc. Am.*, vol. 101, pp. 2299-2310, 1997.
- [5] Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C. and Molholt, G., "The effects of identification training on the identification and production of English vowels by native speakers of Japanese," *Appl. Psycholing.*, vol. 26, pp. 227-247, 2005.
- [6] Bradlow, A. R., "Training non-native language sound patterns: Lessons from training Japanese adults on the English /r/ and /l/ contrast," in *Phonology and Second Language Acquisition*, J. G. Hansen-Edwards and M. L. Zampini, Eds., Philadelphia: John Benjamins Publishing Company, 2008, pp.287-308.
- [7] Iverson, P., & Evans, B.G., Learning English vowels with different first language vowel systems II: Auditory training for native Spanish and German speakers. *J. Acoust. Soc. Am.*, vol. 126, pp. 866-877, 2009.
- [8] Lively, S. E., Pisoni, D. B., Yamada, R. A., Tokhura, Y., and Yamada, T., "Training Japanese listeners to identify English /r/ and /l/ III. Long-term retention of new phonetic categories," *J. Acoust. Soc. Am.*, vol. 96, pp. 2076-2087, 1994.
- [9] Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *J. Acoust. Soc. Am.*, vol. 89, pp. 874-886, 1991.
- [10] Nishi, K., & Kewley-Port, D., Training Japanese listeners to perceive American English vowels: Influence of training sets. *J. of Speech, Lang., and Hear. Research*, vol. 50, pp. 1496-1509, 2007.
- [11] Nishi, K., & Kewley-Port, D., Nonnative speech perception training using vowel subsets: Effects of vowels in sets and order of training. *J. of Speech, Lang., and Hear. Research*, vol. 56, pp. 1480-1493, 2008.
- [12] Strange, W., and Dittmann, S., "Effects of discrimination training on the perception of /r/-/l/ by Japanese adults learning English," *Percept. and Psychophy.*, vol. 36, pp. 131-145, 1984.
- [13] Wong, W. S., "Training the Perception and Production of English Vowels /e/ and /æ/ by Cantonese-speaking Secondary School Students." MPhil dissertation, Dept. Eng., The Chinese Univ. of Hong Kong, Hong Kong, 2010.
- [14] Wong, W. S. "Training the Perception and Production of English Vowels /i/-/i:/, /e/-/æ/ and /o/-/u:/ by Cantonese ESL Learners in Hong Kong," PhD dissertation, Dept. Eng., The Chinese Univ. of Hong Kong, Hong Kong, 2013.
- [15] Wong, J. W. S., "Training the Perception and Production of English /e/ and /æ/ of Cantonese ESL Learners: A Comparison of Low vs. High Variability Phonetic Training." Proceedings of the 14<sup>th</sup> Australasian International Conference on Speech Science and Technology (SST 2012). Sydney, Australia, pp. 37-40., 2012.
- [16] Iverson, P; Hazan, V; Bannister, K; (2005) "Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults." *J. Acoust. Soc. Am.*, vol. 118 (5) 3267 - 3278.
- [17] Ikuma, Y., Akahane-Yamada, R., "Effects of learners' L2 proficiency and acoustic/semantic contexts of stimuli on the phoneme identification training." *J. Acoust. Soc. Am.*, vol. 120, p. 3173, 2006.
- [18] Chan, A. Y. W., & Li, D. C. S. English and Cantonese Phonology in Contrast: Explaining Cantonese ESL Learners' English Pronunciation Problems. *Language Culture and Curriculum*, vol. 13, pp.67-85, 2000.
- [19] Hung, T. T. N., "Towards a Phonology of Hong Kong," in *Hong Kong English: Autonomy and Creativity*, K. Bolton, Ed., Hong Kong: Hong Kong University Press, pp. 119-140, 2005.
- [20] Meng, H., Zee, E., & Lee W. S., A Contrastive Phonetic Study between Cantonese and English to Predict Salient Mispronunciations by Cantonese Learners of English. Unpublished article. The Chinese University of Hong Kong, 2007.
- [21] Hong Kong Examinations and Assessment Authority (2014, Jan 6). *Development of Public Exams in Hong Kong*. [Online]. Available: [http://www.hkeaa.edu.hk/en/recognition/develop\\_hk\\_pub\\_exam/hkcee/](http://www.hkeaa.edu.hk/en/recognition/develop_hk_pub_exam/hkcee/).