A new perspective of CALL software for English perceptual training in pronunciation instruction

Fei-Hsuan Liao

Chienkuo Technology University, Taiwan, ROC
fhliao@ctu.edu.tw

Regular Paper

In this research, the effect of CALL software on English perceptual training is investigated. English Perceptual Pronunciation Training software is designed to train the learner's perception of English vowels by building up adequate acoustic images in the learner's mind. Non-English majors are selected as experiment subjects, who receive English perceptual training via CALL software and subsequently take two distinctive tests: Discrimination and Identification tests. Two different questionnaires are distributed among subjects before and after the instruction: a questionnaire enquiring about their English learning experience and an evaluation questionnaire constructed particularly from the learner's viewpoint. Experimental results are discussed, and the influence of subjects' English-language learning experience is taken into account. Conclusions will be drawn in relation to the tutorial design of this software and the possibilities and perspectives explored by CALL software for integrating perceptual training in pronunciation instruction.

Introduction

In pronunciation instruction, pronunciation perceptual training has been neglected for some time and, as we have experienced in the language classroom, it has not been formally incorporated into any pronunciation training program, at least to the knowledge of the researcher. Moreover, in second/foreign language learning, a great majority of native speakers of Mandarin are still struggling with English pronunciation (Molholt, 1988; Wang, 1988; Chen, Robb, Gilbert, & Lerman, 2001) despite having made tremendous efforts, especially for

those who have chronically disadvantageous English learning experiences. This being the case, it is one of the research objectives to ascertain whether we can find an alternative way to help the learner grasp the main characteristics of English sounds. Taking advantage of modern technology, the effect of newly developed software on perceptual training is investigated. It is assumed that the software is able to revolutionize the traditional pronunciation training program where articulatory training dominates classroom instruction, and contribute to integrating perceptual training with articulatory training in pronunciation instruction.

Literature review

As is predominantly practiced in many language classrooms, English pronunciation instruction puts much emphasis on the aspect of articulatory training. Perceptual training, on the contrary, is always excluded from normal training programs, or at best, only occupies a small part in each program. Since repeated procedures of articulatory training do not guarantee a better performance from our language learners and failure to achieve accuracy reoccurs, reconceptualization of the status of perceptual training may explore a new perspective for pronunciation instruction.

Perceptual training aims to train the language learner to identify target sounds or segments by means of phonetic contrasts (Shimamune & Smith, 1995). It was assumed that, once the learner receives the auditory information, which can constitute a cognitive perception and reinforce the acoustic image, the learner can capture the phonetic characteristics of target sounds. As Leather pointed out, the danger of prior articulatory training exists in the sensory feedback from the learner's own speech because the production of sounds at the initial stage may be detrimental in nature and cause an inadequate acoustic image in the mind (Neufeld and Schneiderman, 1980; Leather, 1983). In line with this reasoning, sustaining argumentation was proposed and empirical investigation conducted. To emphasize the priority of ear-training, Odisho encouraged training in perception and recognition of new sounds and regarded perceptual training as an essential part of pronunciation instruction (Odisho, 1992). Empirical research showed that a subject receiving listening instruction but no pronunciation training improved his pronunciation to a considerable extent, and therefore it was concluded that the ability of listening discrimination actually enhanced language pronunciation (Shimamune & Smith, 1995). To be more specific, Gimson indicated that efforts should be directed to cultivate the ability to distinguish phonemic characteristics of sounds between the native language and the target language (Gimson, 1989). This idea proved to be far from easy, especially when the learner is liable to be confronted with similar, but not identical, L1 sounds (Flege, 1987; Chen, Robb, Gilbert, & Lerman, 2001). Simply put, evidence suggested that an appropriate pronunciation training program needs to include perceptual training as a regular practice.

A controversial argument was brought forth that perceptual training should precede articulatory training in that, as language competence comes before language performance, so should phonemic perception precede language production (Gilbert, 1987; Flege, 1987; Neufeld & Schneiderman, 1980; Leather, 1983; Pennington & Richards, 1986). Although earlier researchers tended to be cautious that the association between perception and articulation is actually a complicated issue, a few recent empirical studies researched factors that could affect the perception of sounds in foreign languages and argued for the relationship between language perception and production (Bohn & Flege, 1990; Flege, Bohn &

Jang, 1997). In their study, Flege, Bohn and Jang found that the accuracy of English vowel perception and production was a function of subjects' experience of English language, and implicitly proved that the perception and production of target sounds in non-native languages were in a way related (Flege, Bohn & Jang, 1997).

The employment of modern technology in language classroom has become a trend since the 1980s and it is accepted that **CALL** allows second/foreign language users more liberty in learning, creates a more individualized learning environment, and thus fosters autonomous learning (Molholt, 1988; Swann, 1992; Jimenez & Perez, 2002; Wagener, 2006; Figura & Jarvis, 2007; Ma, 2007; Fischer, 2007). The benefits of **CALL** are much agreed upon in terms of its potential to strengthen learning motivation (Chang, 2005) and provide immediate feedback that encourages subsequent learning (Heift & Rimrott, 2008; Hmard, 2006). It is therefore claimed that **CALL** is able to renovate traditional language classroom practice and provide more efficient instruction in pronunciation training (Pennington, 1999). An extensive review of the literature, however, reveals that only a limited amount of research investigates the effect of perceptual training for English vowels via computer-based instruction (Lambacher, 1999; Hazan, Sennema, Iba, & Faulkner, 2005). Research where acoustic data were used to provide visual aid for Japanese learners in order to improve their perception and production of English consonants were proved to be effective (Lambacher, 1999). Similarly, when audio vs. audiovisual media were employed for perceptual training of English consonants, it was found that the audiovisual presentation was superior, and additionally the improvement of pronunciation corresponded to perceptual enhancement (Hazan, Sennema, Iba, & Faulkner, 2005). More research relevant to the present study was conducted by Wang and Munro (2004). By drawing language learners' attention to vowel quality, rather than to vowel length, they selected three sets of vowel contrast for perceptual training. Synthetic word pairs were generated as experimental stimuli and learners of Mandarin and Cantonese speakers chosen as participants. The results of identification test showed that significant difference was obtained between pre- and post-testing for the experimental group, but not for the control group, and that a retention test three months later also presented a similar difference between the pre-test and retention test for the same group (Wang & Munro, 2004). This research is of great importance because it indisputably illustrates the effectiveness of a computer-assisted instruction in a pronunciation training program.

As education professionals often worry, it has been warned that the design of **CALL** software does not take into consideration the aspect of pedagogy, only purely technological aspects. A more comprehensive evaluation for multimedia development is therefore suggested to involve pedagogic requirements (Warschauer & Healey, 1998; Pennington, 1999). Apart from the emphasis placed on perceptual training in contrast to articulatory training, the novelty of this research originates in the fulfillment of such a requirement. A questionnaire is constructed from the learner's point of view, which is mainly concerned with the perceived usefulness, effectiveness and user-friendliness of the to-be-attested software. It can be argued that asymmetry exists that, although the role of learners has long been recognized as influential for successful language learning, their attitudes and opinions have not been valued as much in the design of **CALL** software (Scholfield & Ypsiladis, 1994; Bader, 2000; Kessler & Plakans, 2001; Lasagabaster & Sierre, 2003; Murray & Barnes, 1998). For this reason, one aim of this research is to highlight the role of the language learner in the development of **CALL** software.

Multiple goals are conceived in this research. First, the effect of **CALL** software intended **87**

to renovate the content of a current pronunciation instruction program is investigated. Second, understanding the learning experience of Taiwanese students may provide insightful clues for the development of future pronunciation courses. Third, a reiteration of the effectiveness of **CALL** software in the language classroom may also be expected. Fourth, the voice of **CALL** software users needs to be recognized and thus, in the discipline of language learning, the language learner holds the key to the evolution of **CALL** software.

Two research hypotheses are assumed. First, it is expected that the use of **CALL** software will produce significant differences in two experimental Discrimination and Identification tests. In addition, it is expected that participants' English proficiency and the experience of prior pronunciation courses will affect their performance on both tests.

Methodology

Participants

Three classes of non-English majors in a technology university were chosen as participants for this experiment. In total, 131 students took part in this experiment, but only 123 questionnaires were counted valid. The experiment was conducted in a "Practical English" course, which continued for two semesters and was instructed by the same language teacher for two hours a week. An important pedagogic reason for selecting the software as one part of lessons is that the textbook used in this course included pronunciation instruction on English consonant contrasts, but no vowel contrasts were provided, and this software could therefore provide adequate supplementary material for this course. Note that, since the "Practical English" course was scheduled for only one year, it was decided that, in order to provide the most pedagogic benefits for all students, an ideal research design of control vs. experimental groups would be discarded. Instead, the one-group pre-test/post-test design was adopted. In other words, all participants would be trained with the software and their performance before and after the training would be measured and compared.

Apparatus

The software English Pronunciation Perceptual Training Program (**EPPT**) is designed to train the perception of English vowels as an initial step to perceptual training because this set of sounds is more resonant and the phonemic contrasts are confusing. It consists of four parts: function, tutorial lessons, games, and self-assessment. Classroom instruction is centered on the tutorial lessons of four sections: phonemes, minimal pairs, Chinese vs. English, and degree of nativeness. Section one, "phonemes," introduces individual vowels along with word examples. The second and third sections are parallel in that, as a comparison is made between similar English phonemes in "minimal pairs," so it is made between Chinese and English in "Chinese vs. English." The last section, "degree of nativeness," demonstrates three pronunciations of various degrees of "nativeness," among which only one is standard and highly recommended to learners. After a brief introduction of the software, a 6-hour training session in a language lab is given and participants can operate the software at their own pace with the assistance of the language teacher.

Procedure

Pre- and post-experimental questionnaires. Before the implementation of this experiment, participants were given questionnaires about their English-language learning background to better understand variables which could possibly affect experimental results. In order to ensure that participants could fully understand all questions and feel comfortable in providing reliable data, the questionnaire was written in their mother tongue (i.e. Mandarin). Altogether eleven questions were included (Appendix A: English version of the questionnaire).

A post-experiment questionnaire was also constructed, which enquired about the perceived usefulness, effectiveness, and user-friendliness of the software after participants completed the training (Appendix B). In particular, this was designed from the viewpoint of the user, (i.e., the language learner). To the knowledge of the researcher, few studies evaluated CALL software of language learning from the learner's perspective and consequently lacked an important link in the development of CALL software (Scholfield & Ypsiladis, 1994; Kessler & Plakans, 2001; Lasagabaster & Sierre, 2003). Although the role of the learner has long been regarded essential for successful language learning, the language learner's attitude and opinion has not been valued as much. Unlike the conventional evaluative procedure executed by software experts or language teachers, this questionnaire accentuated the learner's viewpoint for software evaluation.

Six fundamental sections were designed, each containing various numbers of questions, including 1) general overview, 2) interface design, 3) operational efficiency, 4) efficacy of the software, 5) usefulness of self-assessment, and 6) ability of vowel contrast. A 5-point Likert scale was adopted so that every question was provided with five choices ranging from 1 to 5: 1) strongly disagree, 2) disagree, 3) no opinion, 4) agree, or 5) strongly agree. To obtain more opinions about the design of this software, the last section of four open-ended questions was offered and participants were allowed to express their opinions more freely.

Experimental tests: Discrimination and Identification tests. Discrimination and Identification pre-tests and post-tests were constructed. Discrimination tests measured the ability to make a correct judgment as to whether two words consecutively pronounced by a native speaker contained the same or different vowels (Appendix C). Both pre-test and post-test contained 20 question items (40 monosyllabic words, which did not appear anywhere in the software). On the other hand, Identification tests required the ability to choose the correct phonemic symbol out of two possible choices after the target word was read by the same native speaker (Appendix D). This kind of test was considered more difficult than the previous one in that participants first needed to recognize a sound and then connect the sound with an appropriate symbol (Wang, & Munro, 2004). Similar to Discrimination tests, Identification tests contained 20 items (20 monosyllabic words, which did not appear in either the EPPT software or Discrimination tests) in pre- and post-tests.

Results

English learning background questionnaires

In this section, we will discuss only those questions which provide more crucial information about the participants' learning experience. Accordingly, only the results of questions 3, **89**

8, and 11 will be presented. However, others will be omitted either because they are highly homogeneous, such as questions 1 and 2, or because they can not be precisely measured, such as questions 4, 5, 6, 7, 9 and 10.

Question 3: Length of exposure to English: In the education system of Taiwan, English courses are obligatory in junior high school and all higher levels. The group of learners in this study were freshmen in a two-year senior college program, meaning that they must have been learning English at least for eight years, (either 3 years for junior high, 3 years for senior high, and 2 years for two-year junior college; or 3 years for junior high and 5 years for five-year junior college). Given the educational background of the participants, the period of eight years was regarded critical in counting the length of English exposure because we could classify the length of exposure to English into two categories: more than 8 years (starting roughly before puberty) and 8 years or less (starting roughly after puberty). As a result, 41% of participants learned English after puberty and 59% before puberty.

Question 8: Experience with prior English pronunciation courses: We assumed that the participants' experience with prior English pronunciation courses might play a role in this experiment. To verify this assumption, question 8 was intended to reveal the participants' experience in prior pronunciation courses and the specific kind of pronunciation courses they had taken. The aggregation of all pronunciation courses showed that more than half of the population had such experience (62.6% to be exact).

Question 11: Experience using **CALL** software: The researcher also expected that experience using **CALL** software could affect the ease with which participants managed the software, especially when they would need to evaluate this software at a later stage. Surprisingly, only seven participants had such experience.

As shown, questions 3 and 8 addressed variables that could potentially affect experimental results and, accordingly, would be treated as independent variables in statistical analysis.

Software evaluation questionnaires

A reliability analysis was also performed and the result is shown in Table 1. As the alpha values of all sections reached 0.8 or higher, we were confident that the questionnaire was reliable (Gay, 1992).

More information could be found in this table. As shown, section 3, regarding operational efficiency, obtained a very high item mean and could indicate that the participants were able to use the software even without much experience in **CALL** software, which been revealed in the English learning background questionnaire. To put it in appropriate terminology, this software was regarded as highly user-friendly. In contrast, section 6 obtained the lowest item mean and had the largest item **SD**, indicating that the participants did not have much confidence in distinguishing similar vowels and that greater variation existed in the ability of participants. More practice would be recommended for this group of learners.

Table 1: Reliability and descriptive analysis for each section of the questionnaire.

Sections	N	Mean	SD	Item mean	α
1) general overview	10	36.99	4.65	3.70	.80
2) interface design	10	37.30	5.37	3.73	.87
3) operational efficiency	12	45.70	5.66	3.81	.86
4) efficacy of the software	13	47.46	7.01	3.65	.88
5) usefulness of self-assessment	10	37.34	5.18	3.73	.88
6) vowel contrast	9	32.30	6.11	3.59	.93

No. of cases: 123

For the open-ended section of the evaluation questionnaire, questions 1 and 2 could be contrasted because they concerned participants' likes and dislikes of the software. 23 participants answered that they liked the section of "tutorial lesson," but only 12 participants indicated that they disliked it. The comparison enabled us to determine that the design of tutorial lessons was successful. Unexpectedly, comparable numbers of participants were found to like and dislike the games part (21 and 19 participants, respectively). We speculated that it was the abundant experience of this generation in using computer games that gave rise to their high expectation for diverse and intriguing computer games.

As for questions 3 and 4, constructive opinions were stated, such as the need to add more word examples for each vowel, the consideration for the randomized order and varied types of test questions in the self-assessment, and finally the addition of more interactive functions, on-line surveys, and animation. Undeniably, these suggestions showed us an explicit direction for improving the software.

Discrimination and Identification tests

To understand the effect of this software, the participants' performance in Discrimination and Identification pre- and post-tests would be our primary data. Paired-sample t-tests were conducted on both Discrimination and Identification tests, with the results shown below. From the following table, we can see that there was a significant difference in the scores for Discrimination pre-test (M=14.32, SD=2.75) and post-test (M=15.17, SD=2.44) conditions, t(122)= −2.935, p <.01 , and similarly for Identification pre-test (M=12.04, SD=2.57) and post-test (M=12.83, SD=2.55) conditions, t(122)=−2.912, p <.01. That is to say, the participants did make progress after using the software, with the difference showing a level of significance at p < .01.

Table 2: Paired-samples t-tests for discrimination and identification tests.

	N	df	Mean	SD	t
Discrimination					
pre-test	123	122	14.32	2.75	−2.935**
post-test	123	122	15.17	2.44	
Identification					
pre-test	123	122	12.04	2.57	−2.912**
post-test	123	122	12.83	2.55	

**p < .01

Prior pronunciation courses, English proficiency, and experimental tests

In the researcher's university, students need to take a university proficiency test every semester. According to the English enforcement policy of this university, three levels are distinguished: level 1 refers to those who score below 30 (out of 100), level 2 to those who score between 31-60, and level 3 to those who are above 60. It was found that for this group of participants, 25 were at level 1, 80 were at level 2, and only 18 were level 3. With the average score being 43.6 out of 100, it was determined that this group of learners was low-intermediate.

Table 3: Numbers of participants, means, and SD for discrimination and identification post-test as a function of proficiency level and experience of prior phonetic course

	Kind of post-tests					
	Discrimination			**Identification**		
	n	M	SD	n	M	SD
With experience of prior phonetic course						
0–30 (Level 1)	17	14.88	2.62	17	12.24	2.25
31–59 (Level 2)	45	15.58	2.04	45	13.56	2.20
60 or above (Level 3)	16	17.00	1.75	16	13.63	2.22
Without experience of prior phonetic course						
0–30 (Level 1)	8	13.50	1.51	8	9.63	1.92
31–59 (Level 2)	30	14.57	2.87	30	12.23	2.91
60 or above (Level 3)	7	13.57	1.62	7	14.00	1.53

Reflection on the results of the English learning background questionnaires led us to assume that both the proficiency level and the experience of prior English pronunciation courses might affect participants' performance on the Discrimination and Identification tests. Therefore, an **ANCOVA** was conducted, treating participants' scores in pretests of both kinds as covariates, and their proficiency level and experience of prior pronunciation courses as independent variables. The result of the descriptive analysis is shown in Table 3. The column "Yes" represents the participants who previously had phonetic-related courses, while "No" refers to those without such experience.

In Table 4, the analysis shows that, while the interaction of these two independent variables did not reach a significant level for the Discrimination post-test, $F(2, 116)=1.361$, $p > .05$), the effect of the experience of prior pronunciation courses did with $F=11.712$, $p < .01$. Similarly, the interaction of these two variables on the Identification tests did not make any significant difference $F(2, 116)=.1.888$, $p > .05$, but the effect of proficiency level did with $F=6.066$, $p < .001$. A post hoc analysis showed that the performance of proficiency levels 3 and 2 were significantly better than level 1, at the levels of $p < .01$ and $p < .05$ respectively.

Table 4: ANCOVA for the effect of prior phonetic course and proficiency level on Discrimination and Identification post-tests.

Source	Discrimination					Identification				
	SS	Df	MS	F	p	SS	Df	MS	F	p
Pre-test (covariate)	20.84	1	20.84	4.11	.045	53.37	1	53.37	10.32	.002
Phonetic courses	53.12	1	53.12	10.46	.002**	19.62	1	19.62	3.79	.054
Proficiency level	24.32	2	12.16	2.40	.096	84.08	2	42.04	8.13	.000***
Phonetic x proficiency	13.82	2	6.91	1.36	.260	19.52	2	9.76	1.89	.156
Error	588.98	116	5.08			599.80	116	5.17		

**p < .01, ***p < .001

These experimental results are noteworthy. For the Discrimination tests, only the variable of prior pronunciation courses produced a significant difference on the participants' performance, while for the Identification tests, only the factor of proficiency level could affect the participants' performance to a significant extent. The different results were expected in that, as mentioned earlier, the Identification tests involved not only the ability to perceive the difference between vowels, but also the ability to recognize the phonemic symbols and therefore were considered more difficult than the Discrimination tests. Viewed in this light, it actually illuminated the distinctive nature of these two kinds of tests.

Discussion

First and foremost, the effect of this software is confirmed by the paired-samples t-tests on both the Discrimination and Identification tests. In other words, the EPPT software does cause a significant difference on participants' performance on both the Discrimination and Identification post-tests. It goes without saying that this research provides an additional piece of evidence for the effectiveness of CALL software in second/foreign language learning. It will be argued that the merit of the software principally lies in the design of tutorial lessons in which three sections deserve our attention: "minimal pairs," "Chinese vs. English," and "degree of nativeness." In the minimal pairs section, participants are presented with mono-syllabic word pairs containing the same sounds but different vowels, and are trained to distinguish the characteristics of the different vowels. The contrast provides an acoustic image that greatly contributes to the perception of English vowels. Second, the "Chinese vs. English" section exceptionally illustrates seemingly similar vowels of Chinese in comparison with English ones and unexpectedly uncovers the Chinese accent. Moreover, in the "degree of nativeness" section, the juxtapositions of careless vs. accurate pronunciations of vowels can enlighten participants to the issue of intelligibility and accuracy in pronunciation and prod them to speak clear English.

Nonetheless, a number of drawbacks cannot be neglected and deserves more discussion. A major one is the restriction of research design. As mentioned in the methodology section, in order to provide most pedagogic benefits for all students, the one-group pre-test-post-test design was adopted, allowing all students to be trained with the software. In so doing, it is not possible to reliably isolate the effect of the software and therefore, we would like to reserve our position in drawing any conclusion. A more careful scrutinizing of the research design reveals that, being more exploratory in nature, the questionnaire for English

learning background is intended to extract variables possibly affecting experimental results, but the researcher finds it difficult to obtain precise information for some questions, such as questions 9 and 10. Future research may modify and redefine these questions. A more intricate and complicated picture emerges while we take more variables into consideration. Two factors, the proficiency level of participants and their experience of prior English pronunciation courses, produce diverse results in the experimental tests. It is shown that while the experience of prior pronunciation courses produces significant difference on the Discrimination post-test, participants' proficiency levels seemingly affects the Identification post-test. Does that mean Discrimination tests are more closely related to the experience of prior pronunciation courses and Identification tests to participants' proficiency levels? Apparently, more evidence is yet to be found.

As far as the technology is concerned, inherent limitations exist. Not until the last two decades have we witnessed the fast development of **CALL** software employed in language teaching. Earlier literature review indicates that perception training has not received much attention in pronunciation instruction programs, let alone efforts spent in the development of favorable software for perceptual training. The **EEPT** software used in this research can be regarded as an original work to explore the possibility of providing adequate perceptual training via modern technology. There is no doubt that this software leaves much room for improvement. Specific suggestions were offered by the participants in their responses to the open-ended section of the software evaluation questionnaire. The concern with the quantity and quality of word examples for each vowel is well stated in that not only a larger amount of word examples is required for each vowel, but they also need to be presented in various contexts to expand the vowel experience of language learners. A recommended re-design of the "self-assessment" section, furthermore, is directed to randomize question order and vary test types. Above all, a modernized view of **CALL** software is presented. For instance, the advice to amplify the interaction between the user and the software and to adopt online service clearly indicates the future trend of **CALL** software development.

Under all these circumstances, we would like to draw a modest conclusion that this software has the potential to improve the perception of English vowels, but for future studies of this kind, a stronger research design is required to build a solid experimental foundation and illustrate benefits. Aside from this, the interference of English proficiency and the experience of prior pronunciation courses was also found to play a role and somewhat compromises the effect of this software. In the future, further investigation should address these issues.

Lastly, on the ground of all the experimental findings, the accomplishment of this research can be summarized as threefold. First, the importance of perceptual training of target languages for second/foreign language learning is highlighted. In the pronunciation instruction program, perceptual training has been neglected for a long time, and this research illustrates the plausibility of incorporating it into a language curriculum and accordingly assigns it a proper place. What is more important, the construction of the software evaluation questionnaire and two vowel perception tests is regarded as highly innovative. The software evaluation questionnaire brings to the foreground the necessity of learners' points of view, which ought to be respected rather than overlooked in the development of effective and user-friendly **CALL** software. For Discrimination and Identification tests, the fact that these two tests produced diverse results means that these two kinds of tests are fundamentally distinctive. While the Discrimination test requires participants to make a distinction between two similar vowels, the Identification test actually demands one more

aspect of cognitive ability, i.e. the knowledge of English phonemic symbols of vowels. The usefulness of these two tests can be attested in more studies. Finally, a strong implication can be drawn that the successful application of software for sound perception allows us to investigate the association between perceptual training and pronunciation performance, and substantiate the assumption that perceptual training holds a key to the success of pronunciation instruction.

References

Bader, M. J. (2000). Choosing **CALL** software: beginning the evaluation process. *TESOL Journal, 9*(2), 18–22.

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.

Chang, M. M. (2005). Applying self-regulated learning strategies in a web-based instruction: an investigation of motivation perception. *Computer Assisted Language Learning, 18*(3), 217–230.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*(6), 427–440.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393–410.

Figura, K & Jarvis, H. (2007). Computer-based materials: a study of learner autonomy and strategies. *System 35*, 448–468.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in **CALL**. *Computer Assisted Language Learning, 20*, 409–442.

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

Gay, L. R. (1992). *Educational research competencies for analysis and application*. New York: Macmillan.

Gilbert, J. (1987). Pronunciation and listening comprehension. In J. Morley (Ed.), *Current perspective on pronunciation: practice anchored in theory* (pp. 29–39). Washington: Teachers of English to Speakers of Other Languages.

Gimson, A. C. (1989). *An introduction to the pronunciation of English*. London: Edward Arnold.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication, 47*, 360–378.

Heift, T. & Rimrott, A. (2008). Learner responses to corrective feedback for spelling errors in **CALL**. *System, 35*, 196–213.

Hmard, D. (2006). Design issues related to the evaluation of learner - computer interaction in a web-based environment: activities vs. tasks. *Computer Assisted Language Learning, 19*, 261–276.

Hong, Y. T. (2004). *Design and evaluation of **CAI** programs of English pronunciation perceptual training*. Unpublished master's thesis, National Chiayi University, Chiayi City, **ROC**.

Jimenez, M. & Perez, J.M. (2002). Learner autonomy and new technologies. *Educational Media International, 39* (1), 61–68.

Kessler, G. & Plakans, L. (2001). Incorporating **ESOL** learners' feedback and usability testing in instructor-developed **CALL** materials. *TESOL Journal*, *10*, 15–20.

Lambacher, S. (1999). A **CALL** tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning, 12*(2), 137–156.

Lasagabaster, D., & Sierra, J. M. (2003). Students' evaluation of **CALL** software programs. *Educational Media International, 40*(3-4), 293–304.

Leather, J. (1983). Second-language pronunciation: Learning & teaching, *Language Teaching, 16*, 198–219.

Ma, Y. L. (2006). *The relationship of learning motivation, learning involving, and learning satisfaction-in the case of EMBA*. Unpublished master dissertation, Yuanze University, Taiwan.

Molholt, G. (1988). Computer-assisted instruction in pronunciation for Chinese speakers of American English. *TESOL Quarterly, 22*(1), 91–111.

Murray, L., & Barnes, A. (1998). Beyond the "wow" factor-evaluating multimedia language learning software from a pedagogical viewpoint. *System, 26*, 249–259.

Neufeld, G. & Schneiderman, E. (1980). Prosodic and articulatory features in adult language learning. In R. C. Scarcells & S. D. Krashen (Eds.), *Research in second language acquisition: selected papers of the Los Angeles second language acquisition research forum* (pp. 163–178). Massachusetts: Newbury House.

Odisho, E.Y. (1992). A comparative study of English and Spanish vowel systems: theoretical and practical implication for teaching pronunciation (**ERIC** Document Reproduction Service No. **ED** 352836). Northeastern Illinois University, Chicago.

Pennington, M. C. (1999). Computer-aided pronunciation pedagogy: promise, limitations, directions. *Computer Assisted Language Learning, 12*(5), 427–440.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*(2), 207–225.

Scholfield, P., & Ypsiladis, G. (1994). Evaluating computer assisted language learning from the learners' point of view. In D. Graddp & J. Swann (Eds.), *Evaluating language: Papers from the annual meeting of the British association for applied linguistics held at the University of Essex, September 1992.* Clevedon, **UK:** British Association for Applied Linguistics in association with Multilingual Matters Ltd.

Shimamune, S., & Smith, S. L. (1995). The relationship between pronunciation and listening discrimination when Japanese natives are learning English. *Journal of Applied Behavior Analysis, 28*, 577–578.

Swann, P. (1992). Computer assisted language learning for English as a foreign language. *Computers Education, 19*(3), 251–266.

Wagener, D. (2006). Promoting independent learning skills using video on digital language laboratories. *Computer Assisted Language Learning, 19*, 279–286.

Wang, C. (1988). *The production and perception of English vowels by native speakers of Mandarin.* Unpublished **MA** thesis, University of Alabama, Birmingham.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*, 539–552.

Warschauer, M., & Healey, D. (1998). Computers and language learning: an overview. *Language Teaching, 31*, 57–71.

Acknowledgements

Author biodata

Fei-Hsuan Liao is an associate professor at Chienkuo Technology University, Taiwan, **ROC**. She has been teaching English as a foreign language for 20 years and obtained a PhD in Applied Linguistics from the University of Durham, **UK**.

Appendix A

Questionnaire for English language learning background

This questionnaire concerns your English language learning background. Please write down related information according to your own experience. What you provide will be considered confidential.

1. Personal information
   Gender: _____ Education system: _____
   Department: _____ Grade: _____
2. Do you have any hearing problems?
   □ Yes
   □ No
3. How many years have you studied English? _____ years.
4. How many hours per week of English course do/did you complete in your school? _____ hours.
5. Apart from these English courses, do you attend additional English courses outside the university?
   □ Yes _____ course, _____ hours/per week
   □ No
6. Do you have constant contact with native English speakers outside English classroom?
   □ Yes, _____ (Nationality)
   □ No
7. Do you regularly spend extra time studying English?
   □ Yes _____ hours
   □ No
8. Have you ever attended English phonetics courses or pronunciation courses?
   □ Yes □ vowels □ consonants □ both □ others, _____ please specify.
   □ No (Please skip questions 9 and 10)
9. If yes, how long ago? ____ years ago.
10. If yes, how many hours in total? ____ hours.
11. Have you ever used **CALL** software? □ Yes. □ No.

Appendix B

Questionnaire for EPPT software

After you have used the **EPPT** software, please answer the questions in this questionnaire. Your opinions and comments help us develop more effective software for English language learning. Seven sections are included, with each containing various numbers of questions. Please indicate that you are, strongly disagree (1), disagree (2), no comments (3), agree (4), or strongly agree (5) by circling your answer. Your patience and assistance are much appreciated.

I · General overview

1. The tutorial section explicitly instructs the distinction between vowels.	1	2	3	4	5
2. The amount of tutorial section is adequate.	1	2	3	4	5
3. The content of the tutorial section masters a proper degree of difficulty.	1	2	3	4	5
4. The elements of the tutorial section can correspond to each other (consistency of structure).	1	2	3	4	5
5. The content of the tutorial section is ordered from simple elements to difficult ones.	1	2	3	4	5
6. The pictures used in the tutorial section are bright.	1	2	3	4	5
7. The words used in the tutorial section are comprehensible.	1	2	3	4	5
8. The recorded sounds in the tutorial section are clear.	1	2	3	4	5
9. The terminology used in the tutorial section is intelligible.	1	2	3	4	5
10. This software can reinforce my motivation to learn English.	1	2	3	4	5

II · Interface design

1. The layout of the software is well organized.	1	2	3	4	5
2. The compositional design of the software is brilliant.	1	2	3	4	5
3. The location of operational buttons is transparent.	1	2	3	4	5
4. The symbols and signs can reflect their functions.	1	2	3	4	5
5. The arrangement of words and background color is consistent.	1	2	3	4	5
6. The fonts used in this software are recognizable.	1	2	3	4	5
7. The amount of words is properly segmented.	1	2	3	4	5
8. The arrangement of words can fit the browsing route of users.	1	2	3	4	5
9. The design of animation is smooth and vivid.	1	2	3	4	5
10. The placement of information is obvious and noticeable.	1	2	3	4	5

II · Operational efficiency

1. The software explicitly explains its objectives.	1 2 3 4 5
2. The software explicitly explains the content range of tutorial section.	1 2 3 4 5
3. The software explicitly explains the operation manner.	1 2 3 4 5
4. The guideline of the software is clear and understandable.	1 2 3 4 5
5. The responding speed of the software is fast.	1 2 3 4 5
6. The software provides solutions to operational errors.	1 2 3 4 5
7. The switching paths between sections are clear and smooth.	1 2 3 4 5
8. The execution operation is precise.	1 2 3 4 5
9. I can control the order of each section.	1 2 3 4 5
10. I can control my own pace.	1 2 3 4 5
11. I can use the software without much practice.	1 2 3 4 5
12. I can quit the software anytime I want.	1 2 3 4 5

IV · Efficacy of the software

1. "Overview" section helps me understand the structure of the software.	1 2 3 4 5
2. "Objectives" section helps me understand the objectives of this software.	1 2 3 4 5
3. "Guidelines" section helps me understand how to operate the software.	1 2 3 4 5
4. The design of male vs. female voices helps me train my perception.	1 2 3 4 5
5. The "phonemes" lesson helps me understand the correct pronunciation of English vowels.	1 2 3 4 5
6. The "minimal pairs" lesson helps me distinguish similar vowels.	1 2 3 4 5
7. The "Chinese vs. English" lesson helps me distinguish between Chinese and English sounds.	1 2 3 4 5
8. The "degree of nativeness" lesson helps me distinguish correct and incorrect pronunciation.	1 2 3 4 5
9. The game "balloon" reinforces my knowledge of phonemic symbols.	1 2 3 4 5
10. The game "beating ghosts" reinforces my knowledge of phonemic symbols.	1 2 3 4 5
11. The "phoneme tests" in the self-assessment part can test my ability of vowel distinction.	1 2 3 4 5
12. The "pronunciation tutor" in the self-assessment part can test my ability of standard and non-standard vowel distinction.	1 2 3 4 5
13. The "minimal pair distinction" can test my ability of similar vowel distinction.	1 2 3 4 5

V · Usefulness of self-assessment

1. The questions corresponds to the content of tutorial section.   1 2 3 4 5
2. The questions are clear and understandable.   1 2 3 4 5
3. The amount of questions is adequate.   1 2 3 4 5
4. The questions control a proper degree of difficulty.   1 2 3 4 5
5. The cues provided for the questions are proper.   1 2 3 4 5
6. The design of questions is varied.   1 2 3 4 5
7. Sufficient time is provided for each question.   1 2 3 4 5
8. The feedback of each question is clear.   1 2 3 4 5
9. The scoring method is reasonable.   1 2 3 4 5
10. The questions help me review tutorial lessons effectively.   1 2 3 4 5

VI · Ability of vowel contrast

1. I can distinguish between /i/ and /ɪ/   1 2 3 4 5
2. I can distinguish between /e/ and /ɛ/   1 2 3 4 5
3. I can distinguish between /ɛ/ and /æ/   1 2 3 4 5
4. I can distinguish between /e/ and /æ/   1 2 3 4 5
5. I can distinguish between /o/ and /ɔ/   1 2 3 4 5
6. I can distinguish between /u/ and /ʊ/   1 2 3 4 5
7. I can distinguish between /a/ and /ʌ/   1 2 3 4 5
8. I can distinguish between /æ/ and /ʌ/   1 2 3 4 5
9. I can distinguish between /æ/ and /a/   1 2 3 4 5

VII · Suggestions
1. What part of the software do you like most?
2. What part of the software do you dislike most?
3. What kind of improvement do you think is needed for this software? Please state it briefly.
   a. Objectives
   b. Guidelines
   c. Tutorial lessons
   d. Games
   e. Self-assessment
   f. Others
4. Is there any additional function you think is necessary for this software?

Appendix C

Discrimination Task (Pretest)

Please indicate whether the vowels on the following pairs are the same or different.

下列各題中你將會聽到一組字，請指出其中的母音是相同或不同。

	Same	Different
1. seat, sit		
2. hood, who'd		
3. made, maid		
4. flap, flop		
5. weak, week		
6. eight, ate		
7. sold, salt		
8. grace, grass		
9. ad, odd		
10. rash, rush		
11. soul, Seoul		
12. ware, wear		
13. aid, ad		
14. clue, clew		
15. bet, bat		
16. cooed, could		
17. rate, rat		
18. hole, whole		
19. coast, cost		
20. lamp, lump		

Appendix D

Identification Task (Pretest)

Please indicate which vowel is pronounced.

請注意聽錄音帶並圈出其中所唸的母音。

1.	itch	/i/	/ɪ/
2.	clue	/u/	/ʊ/
3.	pot	/æ/	/ɑ/
4.	mood	/u/	/ʊ/
5.	nit	/i/	/ɪ/
6.	shake	/e/	/æ/
7.	soot	/u/	/ʊ/
8.	would	/u/	/ʊ/
9.	beat	/i/	/ɪ/
10.	knot	/ɑ/	/ʌ/

11.	lack	/e/	/æ/
12.	sew	/o/	/ɔ/
13.	fun	/æ/	/ʌ/
14.	head	/ɛ/	/æ/
15.	west	/e/	/ɛ/
16.	seek	/i/	/ɪ/
17.	ate	/e/	/ɛ/
18.	hall	/o/	/ɔ/
19.	sock	/æ/	/ɑ/
20.	said	/ɛ/	/æ/