THE JOURNAL OF ASIA TEFL Vol. 6, No. 4, pp. 217-243, Winter 2009

Pedagogical Practice and Students' Perceived Effectiveness of Web-based Automated Speech Evaluation

Mei-jung Wang
National Kaohsiung Hospitality College, Taiwan

Hsueh Chu Chen
Hong Kong Institute of Education, Hong Kong

Automated speech evaluation (ASE) software is designed to provide instant computer-generated feedback. This study examined the pedagogical practices of a Web-based ASE (WASE) utilized in three different ways in college conversation classes and investigated students' perceived effectiveness of WASE in improving their speech. Eighty college students participated in the study. All of the participants were required to practice the lessons provided in the WASE employed in this study, *MyET*, and the scores were taken as formative assessment. A questionnaire was administered at the end of the course to reveal students' perceived effectiveness of WASE in terms of its instructional and user-interface design. The results show that students in the three classes did not perform differently in the lessons provided in *MyET*, including consonants, vowels, words, and sentences, despite the ASE's different kinds of pedagogical practices. Nevertheless, students in Class B and Class C, who practiced pronunciation with *MyET* in class, received similar scores while Class A, who practiced only after class, received significantly lower scores. As for the functions of *MyET*, most of the students held neutral attitudes. Moreover, the students in this study did not find the spectrogram and speech contour displays in *MyET* useful. However, Class C, who received both teacher feedback and ASE

feedback, demonstrated a more favorable attitude toward the ASE than the other two classes.

Key words: web-based learning, automated speech evaluation pedagogical practice

INTRODUCTION

Due to the tight schedule of English classes in Taiwan, sufficient time is not offered for students to speak in the classroom. Even if they have the opportunity to practice pronunciation, assessment is often done by the teachers' intuition (Chen, 2004). Automated speech evaluation (ASE) software seems to compensate for this problem by providing instant computer-generated feedback. The untiring, non-judgmental nature of the computer allows students unlimited opportunities to review any part of the material and receive additional assistance provided by the system (Tsai, 2006).

However, ASE software is still a developing technology that has not yet reached a fully mature stage. One of the limitations of ASE software is that it is designed without a basis in pedagogical theory. For example, the majority of ASE software puts overwhelming emphasis on the decontextualized mechanics of articulation and focuses exclusively on segments rather than prosody (Breitkreutz, Derwing, & Rossiter, 2002; Chen, 2004; Pennington, 1999). In addition, learners do not benefit much from the graphical wave forms presented in software (Chun, 1998; Neri, Cucchiarini, Strik, & Boves, 2002). Many researchers have emphasized the pedagogical effectiveness of using technology in various learning contexts in addition to the effectiveness of the technology innovations (Chun, 1998; Chen & Cheng, 2008; Levy & Stockwell, 2006; Warschauer & Ware, 2006). With many limitations inherent in the design of ASE technology, the particular pedagogical practices that accompany ASE use are critical.

The relationship between technology and pedagogy, as Stockwell (2007)

suggested, can be seen "as a symbiotic one, where they are mutually dependent upon each other, potentially to their benefit, but also potentially to their detriment" (p. 118). Therefore, ASE software needs to be designed to reduce the limitations and to keep up with current pedagogical theories. It has been suggested that more empirical studies should be conducted to investigate the pedagogical benefits of using ASE software in language classrooms (Salaberry, 1996). This study examined the pedagogical practices of an ASE utilized in three different ways (individual use, in-class use with no feedback, and in-class use with feedback) in college conversation classes and investigated students' perceived effectiveness of the ASE in improving their speech.

LITERATURE REVIEW

Traditional ways of teaching pronunciation mainly rely on language instructors' subjective evaluation of learners' speech production and on learners' recognition of the differences between their pronunciation and the native model on the tape (Molholt, 1988). Such precise discrimination and correct production of distinct sounds are believed to serve as essentials for better pronunciation in spontaneous communication (Pennington, 1996). In the current trend of pronunciation teaching, however, the ultimate goal of teaching and learning pronunciation is intelligibility instead of perfect native-like pronunciation (Celce-Murcia et al., 1996; Kenworthy, 1992; Kirkpatrick, Deterding & Wong, 2008; Munro & Derwing, 1995). This section begins with a discussion of the goals of pronunciation training. Feedback in pronunciation training is then reviewed. Issues concerning the effectiveness of corrective feedback and learners' perceptions of ASE are finally examined.

Goals of Pronunciation Training

Traditional pronunciation teaching materials mainly focused on sound

recognition and discrimination, which is one of the beliefs held by many English as a foreign language (EFL) teachers and students. Morley (1991) argued that the learners' goals of pronunciation training should be to achieve a reasonably "intelligible pronunciation" rather than becoming "perfect pronouncers" (p. 500). Instead, the focus should be on functional intelligibility, functional communicability, increased self-confidence, speech-monitoring abilities, and speech modification strategies for use beyond the classroom. These guidelines have been influential for pronunciation training in the classroom and are equally important for a computerized language tutor.

Both Morley (1991) and Murphy (1991) stressed that pronunciation training must focus both on macro- and micro-level perspectives, i.e., that pronunciation training cannot focus solely on the importance of vowel and consonant sounds but must also focus on suprasegmentals (stress, intonation, etc.). In terms of segmental parts, according to Scarcella and Oxford (1994), consonants contain noises that are pronounced with a blockage of some sort of the air passage. Firth (1992) discussed several questions about the consonant difficulties EFL students encounter, for example, consonant omission, insertion, substitution, consonant cluster simplification, and consonants being linked properly in connected speech. Vowels are characterized by a free passage of air. Firth (1992) suggested the following questions should be of concern for teachers: Are the students substituting one vowel sound for another? Does the vowel have the proper length? Are vowels reduced in unstressed syllables? Are vowels being properly linked to other vowels across word boundaries?

As for the suprasegmental parts, Scarcella and Oxford (1994) described the relationship between stress and rhythm: stress contributes to rhythm. Linguists use the term rhythm to refer to the measured movement or musical flow of language. English has a rhythm in which stressed syllables normally occur at regular time intervals. That is why English is often classed a stress-time language. Firth (1992) suggested that teachers should cover the following issues: Can the students use loudness and length to differentiate between stressed and unstressed syllables? Can the students produce appropriate strong and weak stresses? Are content words stressed and

function words unstressed? Can the students link words appropriately within sentences? With respect to intonation, the pitch of the voice with which a voiced sound is pronounced is called its intonation. According to Scarcella and Oxford (1994) and Wong (1993), it conveys and performs grammatical functions in sentences. Brazil, Coulthard, and Johns (1980) pointed out that intonation conveys a speaker's involvement in a conversation as well as a desire to take a turn or leave a conversation. Firth (1992) suggested that teachers should check if the students use the appropriate intonation pattern.

It is clear that most of the current ASE software does not meet these needs. Morley (1991) argued that pronunciation training should be an integral part of communication, and that imitation practice should be used only as necessary and as a short-term component. Instead, some tasks, such as oral reading of a script and small group discussions, should be used to integrate the modified speech patterns into naturally occurring speech. Such exercises are seldom used and difficult, but not impossible, to incorporate into ASE.

Feedback in Pronunciation Training

Traditionally, every error would be corrected immediately, in methods based on behavioral theories of learning, such as audio-lingualism. When this notion of human motivation changed in general in the beginning of the 1970s, so did pronunciation training. For pronunciation training, this led to new methodology stating that not all errors should be corrected and those that are should not be done so immediately, see, e.g., Krashen (1981). One reason is that not all deviations are errors. Another reason is the affective factor that students may lose self-confidence if they are corrected all the time. However, Ancker (2000) found that students still expect, even request, their teacher to correct all their errors. In the survey by Ancker (2000), the vast majority of the teachers (75%) and the majority of teacher trainees (64%) agreed that not every error should be corrected, while the students (76%) thought it should.

Carroll and Swain (1993) investigated different types of corrective feedback in second language learning and found that all groups receiving

feedback, explicit or implicit, performed significantly better at the end of the study than the control group. There is unfortunately no simple measure of the efficiency of different types of corrective feedback, as efficiency is related to both uptake and successful repair. The first difficulty is how the internal process of the students' uptake, i.e., the students have understood the feedback, should be measured and the second that uptake and successful repair, i.e., students actually change the pronunciation, are related, but not in direct correspondence.

El Tatawy (2002) summarized the prerequisites for successful corrective feedback as follows:

1. The feedback should be provided systematically and consistently.
2. The feedback should be clear enough to be perceived as feedback.
3. Time and opportunity for self-repair and modified output should be given.
4. The feedback should be fine-tuned, ie, the teacher's intent, the target error and the learner's perception of the feedback should match closely.
5. The feedback should focus on error at a time and over a period of time in order to be consistent and intensive.
6. The learner's readiness to process the feedback should be accounted for.

With a thorough survey of feedback and language acquisition, Engwall, Bälter, Öster, & Kjellström (2006) added the following prerequisites:

1. The feedback should be adapted to the situation, i.e.[,] different feedback is suitable for different exercises.
2. The learner should be actively involved in pronunciation monitoring and correction.
3. The feedback should promote the learner's communicative self-confidence.

The Effectiveness of Corrective Feedback and Learners' Perceptions of ASE

Pronunciation learning has been considered an essential part in achieving

successful communication, and ASE software serves as an alternative or extra opportunity for learners to practice their speaking skills and receive necessary corrective feedback. Language learners should not only receive a score but should also be given descriptive or analytic feedback on their pronunciation problems (Neri et al., 2002, 2003). ASE software recognizes language learners' speech sounds, detects errors, diagnoses errors, and then provides corrective feedback (Neri et al., 2003).

It remains unclear whether and how language learners can benefit from this kind of ASE software. Some potentials and limitations of an ASE are discussed as follows. In terms of potentials, ASE software has a range of advantages that give the software special promise for language instruction (Anderson-Hsieh, 1992; Chun, 1989; Pennington, 1989, 1996; Pennington & Esling, 1996). First of all, ASE software performs a quick analysis and gives feedback to the user far faster than a human being can. The analysis of a user's speech is also infinitely repeatable, precise, and reliable in the sense of being the same every time. In all these senses, ASE software is superior to human pronunciation teachers. ASE, which does not suffer from limitations of hearing, judgment, or patience, is in many ways more authoritative than human-aided pronunciation instruction.

ASE software can also individualize pronunciation instruction in ways the pronunciation teacher cannot, based on a mechanical analysis of individual student problems and past trials and performance. The computer can, moreover, make available a much wider range of presentations, on demand and on the spot, than a human trainer. In this sense, the computer has the capacity to present both highly individual and highly variable training. In providing learners with a private and individual workspace and various tools, ASE software can help build learners' confidence while developing skills in pronouncing and discriminating sounds and sound patterns of the target language.

Some studies support the above-mentioned arguments. Hardison (2003) found that using audiovisual feedback was a more efficient way to train learners to discriminate contrastive sounds between L1 and L2 than the

audio-only method in a study of Japanese and Korean learning /r/ and /l/ sounds. Eskenazi (1999) suggested that visual display is more effective than oral instruction in terms of prosody. Learners would better understand how to raise or lower their pitch to match the target speech by listening to the native model and watching the visual display (Molholt, 1990). Hardison (2004) further conducted another pre-test-post-test-design study on prosody learning through computer-assisted training with a real-time visual display of the pitch contours. The results revealed a significant effect of improvement in both prosody and segmental accuracy. It is notable here that prosody-focused training resulted in language learners' improvement in segmental learning.

Even though ASE software possesses positive attributes, some limitations remain. For example, certain aspects of pronunciation do not show up well in the visual representations of the speech analysis, such as waveforms, and so cannot generally be trained by such representations. The second limitation is that most of the ASE software is designed for individual use, and most training systems are limited for whole class instruction. Another serious problem is that the emphasis in computer-based work on pronunciation has been toward the decontextualized mechanics of articulation. Few software applications of the technology link mechanical and meaningful dimensions of phonology (Pennington, 1989).

Reeser (2002) indicated that, though learners were highly motivated to use the program to improve their pronunciation with providing of native speech models, it was unclear how visual support helped learners improve their pronunciation, and that the given scores were vague to learners. Zheng (2002) also found that waveform display was of little help for learners as it was difficult for them to modify their speech to match the sample waveform. The pedagogy of the feedback in existing ASE software could also be questioned, as the feedback given is limited to some extent and needs to be interpreted by the user. Hincks (2005) found that beginners using commercial software for practicing English improved, whereas those at the intermediate level did not. The student who spent the most time, 50 hours, with the software actually got slightly worse, most probably due to self-consciousness. This illustrates the

importance of giving feedback that is relevant and interpretable and does not shock the student with too much information.

Therefore, due to such contradictory findings regarding the feedback in ASE software, substantial evidence of the effectiveness of feedback for language learners should be provided through carefully designed empirical studies. The two research questions addressed in this study are as follows:

1. What are the differences in the learners' performances among the three classes with different ASE pedagogical practices, namely, individual use, in-class use with no feedback, and in-class use with feedback?
2. What is the learners' perceived effectiveness of the ASE pedagogy?

METHODOLOGY

The Context

Eighty students participated in the study. They were three classes of students who took the speaking and listening course for General Education English in the first year at a college in Taiwan. There were 27 students in Classes A and C and 26 students in Class B. They were of the same level of English proficiency according to the placement test administered on the first day of the semester. Materials used were the lessons provided by *MyET*.

In a semester of 18 weeks, apart from the first week for orientation and three weeks for tests and holidays, 14 lessons (one lesson for each week) were chosen from *MyET*. For each lesson, the students were requested to record their speech in *MyET*. All of them were required to practice the lessons, and the scores were taken as formative assessment. However, the pedagogical practices were slightly different in the three classes. For Class A, *MyET* was used merely as take-home practice. In other words, the students in Class A did not use *MyET* in class. For Class B and Class C, *MyET* was used both as take-home practice and as in-class practice. That is, the instructor

allocated 20 minutes for the students to use *MyET* in class. They could also go home and continue their practice. The system recorded the latest scores. Aside from the above practice, the instructor gave some feedback according to the performance of the lesson for Class C but did not provide extra practice or individual correction. A questionnaire (see the Appendix) designed by Tsai (2006) was adapted to reveal students' perceived effectiveness of the ASE software in terms of its instructional and user-interface design at the end of the course.

A Brief Introduction to *MyET*

MyET is pronunciation training software featuring an automatic speech analysis system, application of automatic speech recognition (ASR) technology as well as specification of users' pronunciation problems. The functions include recording, tips for articulation, immediate feedback with verbal and visual illustration, multiple options for native speaker models with different accents, Chinese translation, choice of a slower speed to listen again, peer competition, and recording personal performances.

For each lesson, the system scores for the following four independent phonetic analyses: segments, pitch, timing, and emphasis. Scores given by the software are considered valid for learners to measure their progress in improving their speech production according to Chen's (2004) and Lee's (2008) studies.

In the following paragraphs, screenshots of the software are presented to demonstrate how students practiced pronunciation in the ASE software. The design of the interface was selected in the Chinese version to facilitate students' understanding and operation. After logging in to the software, students can choose among tutors with different accents and different modes of practice (sentence practice, role-play, self-evaluation, and Dr. Pimsleur practice), as illustrated in Figures 1 and 2.

FIGURE 1
The Sample Page of Pronunciation Practice

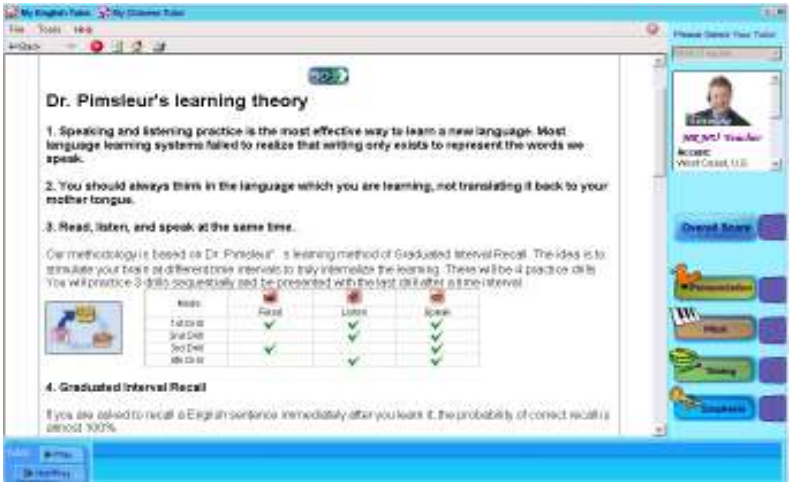

FIGURE 2
Dr. Pimsleur's Learning Theory and Practice Modes

Dr. Pimsleur's learning theory is that speaking and listening practice is the most effective way to learn a new language (http://www.pimsleur-language.com/method.htm). In the language that we are learning, not translating back to our mother language, we should read, listen, and speak at the same time. The methodology of *MyET* is based on Dr. Pimsleur's learning method of Graduated Interval Recall. The idea is to stimulate learners' brains at different time intervals to truly internalize the learning. There are four practice drills. Students were asked to practice three drills sequentially and present with the last drill after a time interval.

After clicking on the sentence, students can listen to it and see the spectrogram, as shown in Figure 3.

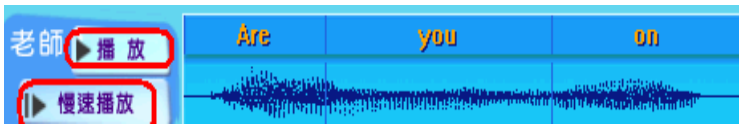

FIGURE 3
The Spectrogram

Students can record their pronunciation and compare their spectrograms with that of the tutor's, as shown in Figure 4.

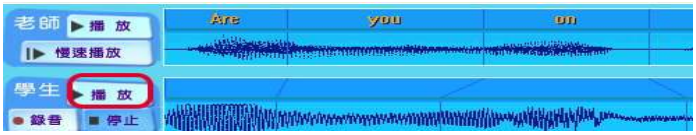

FIGURE 4
Comparison of the Spectrograms

To see the result, students can click on the bottom on the list, and they will see the scores and comments in each category, as shown in Figure 5.

FIGURE 5
The Scores and Comments in Each Category

Feedback on pronunciation was a descriptive account of their poorly or wrongly produced sounds in Figure 6, along with a 3D animation of a mouth in Figure 7. Feedback on intonation, fluency, and volume is similar to that on pronunciation.

FIGURE 6
Feedback on Pronunciation

FIGURE 7
3D Animation of a Mouth

RESULTS

The following sections are preliminary analyses of the data. Students' scores for weekly practices of the three classes are presented in Table 1. As Table 1 shows, the three classes demonstrated no differences in the performance of consonants, vowels, words, and sentences. However, students in Class B and Class C received similar scores while Class A received significantly lower scores (F=7.56; p<0.02). A possible explanation is that the students in the class have better concentration on the tasks than those who practice at home. It is also worth further investigation to see if the equipment plays a role in the process of scoring.

TABLE 1
Comparisons of the Weekly Practices of the Three Classes

	Total	Segments	Pitch	Timing	Emphasis
Consonants					
Class A	83.28	78.03	88.56	88.56	87.79
Class B	88.82	83.51	92.29	96.71	92.27
Class C	88.47	83.23	92.85	96.24	91.21
Vowels					
Class A	86.52	81.66	94.10	86.22	95.82
Class B	88.84	84.37	95.23	89.18	98.23
Class C	89.40	85.10	95.91	89.24	98.25
Words					
Class A	83.28	78.03	88.56	88.56	87.79
Class B	88.82	83.51	92.29	96.71	92.27
Class C	88.47	83.23	92.85	96.24	91.21
Sentences					
Class A	81.98	83.93	75.61	85.50	76.85
Class B	85.60	85.51	83.69	88.28	83.90
Class C	84.09	83.50	81.87	88.00	82.94

Note: The scoring range is between 0 and 100.

The results of the questionnaire are shown in Table 2, including the mean scores and standard deviations of all the items in Question 1: overall score, four independent phonetic analyses (segmental, pitch, timing, and emphasis), and the recording, and (re)playing functions that the *MyET* program provides. As Table 2 shows, most of the students held neutral attitudes toward the functions of *MyET*. The top three functions that were considered helpful are comments on the segment analysis, comments on timing, and comments on emphasis. Meanwhile, the 3D animation of a mouth and the contrast display of the spectrogram with model utterance were regarded as functions that were of little help. It seems that the students in this study prefer feedback with words instead of figures. Maybe it is due to the fact that students could not improve their pronunciation by simply looking at the spectrograms and the animation. This finding is the same as that of Tsai's study (2006).

As for the differences among the three classes (Table 3), the students in Class A considered the pitch analysis and replaying function helpful, while

Class B regarded segment analysis and recording as the two most helpful functions. As for Class C, pitch analysis and overall scores were the top two most helpful functions. No conclusions could be drawn here because different classes favor different functions.

TABLE 2
Mean Scores on the Effectiveness of Items in Question 1

	All Mean	All SD
Overall score	3.14	.12
Segment analysis	3.05	.31
Phonetic symbol display and scores	3.00	.47
Tips for pronunciation of segments	2.10	.58
3D animation of mouth	2.56	.27
Contrast display of spectrogram with model utterance	2.93	.39
Comment	3.34	.41
Pitch analysis	3.04	.44
Display of syllable and key	2.99	.40
Contrast display of intonation with model utterance	3.03	.41
Comment	3.08	.27
Timing	2.97	.33
Display of syllable and speed	2.91	.47
Comment	3.25	.23
Emphasis	3.00	.37
Display of emphasis	2.85	.31
Comment	3.20	.07
Recording function	3.15	.18
(Re)play function: click on certain regions of the waveform or the phonetic symbols for pronunciation of the target syllable or segment	3.14	.12

Note: 5 point scale (1 = not helpful, 5 = very helpful)

As for Question 2, by studying the acoustic spectrum by themselves, students in Classes A and B thought that they were not able to understand the weak points in their pronunciation. In contrast, Class C expressed that the teacher's feedback could help improve their pronunciation. Regarding Questions 3 and 4, most of the students believed that they could become aware of the correctness (or incorrectness) of their own pronunciation after

practicing several times with the assistance of the program.

TABLE 3
Mean Scores on the Effectiveness of Items in Question 1 for Three Classes

	CA	**CA**	CB	CB	CC	CC
	Mean	SD	Mean	SD	Mean	SD
Overall score	3.07	1.174	2.82	.863	3.50	.707
Segment analysis	3.07	1.141	3.07	.979	3.28	.895
Phonetic symbol display and scores	3.11	1.219	2.71	.810	3.33	.907
Tips for pronunciation of segments	3.07	1.072	2.50	.839	3.44	.705
3D animation of mouth	2.78	1.219	1.71	.854	1.83	.985
Contrast display of spectrogram with model utterance	2.67	.877	2.25	.752	2.78	1.166
Comment	2.74	1.130	2.68	.945	3.39	.916
Pitch analysis	3.30	1.068	2.96	.922	3.78	.808
Display of syllable and key	3.37	1.043	2.54	.744	3.22	.808
Contrast display of intonation with model utterance	3.11	.974	2.54	.999	3.33	.840
Comment	3.04	1.091	2.61	.737	3.44	.856
Timing	2.96	1.055	2.89	.916	3.39	.979
Display of syllable and speed	3.04	.898	2.61	.875	3.28	.958
Comment	2.81	1.001	2.50	.638	3.44	.984
Emphasis	3.22	1.013	3.04	.881	3.50	.707
Display of emphasis	3.15	1.064	2.57	.790	3.28	.958
Comment	2.85	.989	2.54	.838	3.17	.924
Recording function	3.19	.962	3.14	.970	3.28	1.074
(Re)play function: click on certain regions of the waveform or the phonetic symbols for pronunciation of the target syllable or segment	3.33	1.144	2.96	1.071	3.17	1.150

Note: 1. 5 point scale (1 = not helpful, 5 = very helpful)
2. CA= Class A, CB= Class B, and CC=Class C

When it comes to the comments on or an overall evaluation of this software program, students' responses revealed that the ASE software had the following advantages:

(1) It is helpful for improving pronunciation;
(2) It is good for self-study; and

(3) It is fun.

Most of the students expressed that the pronunciation of the model was of great help. By listening to and imitating the native speakers' pronunciation, the students improved their own pronunciation. Moreover, the web-based AES allowed them to practice at their own pace after class. In addition, the students found it amazing that they could practice pronunciation with computers.

However, negative comments were classified as follows:

(1) The recording function is not good;
(2) The scoring criteria are not valid; and
(3) It is not user-friendly.

Sometimes the students pronounced some words or sentences several times, but the recording didn't work. As a result, the students had to adjust their recording again and again. This process was depressing. In addition, they doubted the validity of the scoring criteria because some students found that they unintentionally spoke Chinese and received scores. Finally, the system was not user-friendly. Some students were tired of the instability of the system.

Generally speaking, most of the students remarked that the ASE software was helpful for their pronunciation, especially for self-study. In addition, it is fun to have such an ASE learning experience. However, students felt frustrated when they failed to record their sounds. Moreover, students complained about the validity of the scores. Students wondered why they received scores even when they hadn't finished or when they spoke Chinese.

CONCLUSIONS AND SUGGESTIONS

Despite the different kinds of pedagogical practices of the ASE software,

students in the three classes did not perform differently in the lessons provided in *MyET*, including consonants, vowels, words, and sentences. Nevertheless, students in Class B and Class C, who practiced pronunciation with *MyET* in class, received similar scores while Class A, who practiced after class, received significantly lower scores. As for the functions of *MyET*, most of the students held neutral attitudes. Moreover, the students in this study did not find *MyET*'s spectrogram and speech contour displays useful. However, Class C, with the availability of both teacher feedback and ASE feedback, demonstrated a more favorable attitude toward the ASE than the other two classes.

As a whole, most of the students remarked that they benefited from the ASE software when they practiced pronunciation by themselves. In addition, they found it interesting to experience the functions provided by an ASE. However, students felt frustrated when the recording went wrong. Moreover, they complained about the validity of the scores. They hoped that the scoring system could be validated to a great extent.

One of the limitations is that the treatment of the three groups seemed unfair, because students in Class C worked more and received more instruction than those in Class B and Class A, so the results were quite obvious regardless of other differences. The other limitation is that no human scoring was provided to compare the validity of scoring in the ASE. In future studies, more equal treatment for different groups such as giving more appropriate guidance and equivalent feedback would interesting to examine.

Although the findings of the study suggested some forms of feedback provide learners with salient improvement in pronunciation, some limitations still remain, and further studies are recommended. First of all, the time available for training was limited. As the experiment was held in a conversation course, pronunciation training was part of the course components. In terms of pronunciation learning, the in-class practice time was relatively short. For the take-home practice group, we found that it was not easy to track or quantify their practice time after home. The duration of the training effect was also not examined. Although Classes B and C showed

significant improvement, it was unclear whether the learners could maintain the learning effect. Furthermore, the types and quality of feedback from computers and teachers are worth investigating. In this study, a complete lesson plan for teacher feedback was not rendered. It was suggested that a similar research design with corrective feedback procedures would be able to verify whether computer feedback or teacher feedback indeed works for language learners. Finally, descriptions of second language learners' phonological acquisition (e.g., consonants, vowels, stress, and intonation) can be explored in great detail. With these analyses, learners' incidental and developmental errors could be identified. Further remedial treatment and instruction could be supplied. Consequently, the results generated from these future studies could validate the effectiveness of different types of feedback with more substantial evidence.

In light of the findings of the current study, some pedagogical suggestions for web-based pronunciation learning are presented. The power of the new computer technologies-pronunciation ASR tutor, such as *MyET*, to some extent has been validated, providing learners with more opportunities to produce the target language and to have extensive interaction with computers. Instructional materials supported by ASR could be an effective alternative for teaching and learning about speaking and pronunciation. ASR software serves as a private tutor for speaking practice and attends to learners' affective domains because it allows learners do repetitive trials. Learners have individual attention and do not need to compete with their peers. They can learn to communicate in a less threatening environment and get feedback from computers quickly. Learners can have better control of their learning pace and might have enhanced their confidence. Combined with a well-designed tracking mechanism, the feedback provided by the recognizer could be a valuable record for learners' later reflection and further practice. It thus helps relieve teachers' loads.

Additionally, learners can hear the speech models provided by native speakers with different accents. In terms of world English, the explicit teaching of phonemes, stress, and intonation with different inner cycle

accents (e.g., General American (GA), Received pronunciation (RP)) via the computer is effective in facilitating the development of communicative competence. During the communication process, understanding of different foreign accents as well as use of non-foreign-sounding accents could play a key role in successful communication. For effective communication, learners should enrich their phonological knowledge at both segmental and supra-segmental levels to improve intelligibility.

As claimed by Greenspan and Lewis (2002), language activities should be motivating enough to fuel learners' desire to develop receptive and expressive language. In addition to the innovational functions of ASE software, pedagogical practice of such an instrument plays an important role in the adoption of technology. Chen and Cheng (2008) emphasized that it is vital to conduct more classroom-based studies with automatic feedback to probe more deeply into the teaching and learning processes and measure the effects of automatic feedback provided by technology. Moreover, while the design of ASE needs improving, students still need corrective feedback and encouragement from teachers or their peers. In other words, human facilitation should be available in the learning environment. Consequently, this requires a sound pedagogical framework for teachers to integrate technology into speaking and listening courses and create a learning environment with appropriate learning tasks and support.

THE AUTHORS

Mei-jung Wang is assistant professor in the Dept. of Applied English of National Kaohsiung Hospitality College in Taiwan. Her current research interests cover web-based learning and ESP. Her recent publications in 2009 include ‘Web based projects enhancing English language and generic skills development for Asian hospitality industry students’ (*Australasian Journal of Educational Technology)* and ‘The ability of college students to think critically as displayed in online discussion’ (*Taiwan Journal of TESOL*).

Email: sebrina@mail.nkhc.edu.tw

Hsueh Chu Chen is an assistant professor of English at the Hong Kong Institute of Education. Her chief areas of interest are second language phonology and experimental phonetics.
Email: hsuehchu@ied.edu.hk

REFERENCES

Ancker, W. (2000). Errors and corrective feedback: Updated theory and classroom practice. *FORUM, 38,* 20-26.

Anderson-Hsieh, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System, 20*, 51-62.

Brazil, D., Coulthard, M., & Johns, C. (1980). *Discourse, intonation, and language teaching.* London: Longman.

Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2002). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*, 51-61.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generations. *Studies in Second Language Acquisition, 15,* 357-386.

Celce-Murcia, M., & Goodwin, J. (1991). Teaching pronunciation. In M. Celce-Murcia (Ed.), *Teaching English as a second language* (pp. 136-153). NY: Heinle and Heinle.

Chapelle, C. A. (2003). *English language learning and technology.* Amsterdam: John Benjamins.

Chen, H. J. (2004). Automatic speech recognition and oral proficiency assessment. *Proceedings of International Conference on English Language Teaching Instruction and Assessment 2004*, pp. 85-102. Taiwan: National Chung Cheng University.

Chen, K. (2004). *An Investigation on the Impact of ASR Software Feedback on EFL College Students' Pronunciation Learning*. Unpublished master's thesis. National Taiwan Normal University, Taiwan.

Chen, C. F., & Cheng, W. Y. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology, 12*(2), 94-112.

Chun, D. M. (1989). Teaching tone and intonation with microcomputer. *CALICO Journal, 7,* 21-46.

Chun, D. M. (1998). Signal analysis software for teaching discourse intonation. *Language Learning and Technology, 2*(1), 61-77. As retrieved on 10 Aug 2006 from http://llt.msu.edu/vol2num1/article4/index.html.

El Tatawy, M. (2002). Corrective feedback in second language acquisition. Working feedback in second language acquisition. *Working papers in TESOL & Applied Linguistics Teacher's College, Columbia University.* As retrieved on 25 April, 2009 from http://journals.tc-library.org/index.php/tesol/article/view/160/158.

Engwall, O., & Balter, O. (2007). Pronunciation Feedback from Real and Virtual Language Teachers. *Computer Assisted Language Learning, 20*(3), 235-262.

Eskenazi, M. (1999). Using automatic speech processing for foreign language pronunciation tutoring: some issues and a prototype. *Language Learning & Technology, 2*(2), 62-76.

Engwall, O., & Bälter, O. (2007). Pronunciation feedback from real and virtual language teachers. *Computer Assisted Language Learning, 20*(3), 235-262.

Firth, S. (1992). Pronunciation syllabus Design: A question of Focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 173-183). Oxford: Oxford UP.

Greenspan, S. I., & Lewis, D. (2002). *The affect-based language curriculum (ABLC): An intensive program for families, therapists, and teachers.* Bethesda, MD: Interdisciplinary Council on Developmental and Learning Disorders.

Hardison, D. M. (2003). Acquisition of second-language speech: effects of visual cues, context, and talker variability. *Applied Psycholinguistics, 24*(4), 495-522.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: quantitative and qualitative findings. *Language Learning & Technology, 8*(1), 34-52.

Hincks, R. (2005). *Computer support for learners of spoken English*. Unpublished doctoral dissertation. KTH(?), Stockholm, Sweden.

Kenworthy, J. (1992). *Teaching English pronunciation*. New York: Longman Publishing.

Kirkpatrick, A, Deterding D., Jennie. (2008). The international intelligibility of Hong Kong English. *World Englishes, 27*(3), 359-377

Krashen S. (1981). *Principles and Practice in Second Language Acquisition. English Langauge Teaching Series.* London: Prentice-Hall.

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer-assisted language learning*. Mahwah, NJ: Lawrence Erlbaum.

Lee, S. T. (2008). *Teaching pronunciation of English using computer assisted*

learning software: An action research study in an institute of technology in Taiwan. Unpublished doctoral dissertation. University of Reading, England.

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

Molholt, G. (1990). Spectrographic analysis and patterns in pronunciation. *Computers and the Humanities, 24,* 81-92.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, *25*, 481-520.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 25, 5*1-75.

Neri, A., Cucchiarini, C., & Strik, W. (2003). *Automatic speech recognition for second language learning: How and why it actually works.* 15th ICPhs Barcelona.

Neri, A. Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning, 15*, 441-467.

Pennington, M. C. (1989). Application of computers in the development of speaking and listening proficiency In M.C. Pennington (Ed.) *Teaching languages with computers: The State of the Art* (pp. 97-122). Houston, TX: Athelstan.

Pennington, M. C. (1996). *Phonology in English language teaching: An international approach.* London: Longman.

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

Pennington, M. C., & Esling, J. H. (1996). Computer-assisted development of spoken language skills, In M.C. Pennington (Ed.), *The Power of CALL* (pp. 153-189). Houston, TX: Athelstan.

Reeser, T. (2002). "Tell me more French," Software review. *CALICO Journal, 19,* 419-428.

Salaberry, M. R. (1996). A theoretical foundation for the development of pedagogical tasks in computer mediated communication. *CALICO Journal, 14*(1). 5-34.

Scarecella, R. & Oxford. R. L. (1994). Second Language Pronunciation: State of the Art in Instruction. *System, 22*(2), 221-230.

Stevick, E. W. (1978). Toward a practical philosophy of pronunciation: Another view. *TESOL Quarterly, 12,* 145-150.

Stockwell, G. (2007). A review of technology choice for teaching language skills in the CALL literature. *ReCALL, 19*(2), 105-120.

Tsai, P. H. (2006). Bridging pedagogy and technology: User evaluation of pronunciation oriented CALL software. *Australasian Journal of Educational Technology*, *22*(3), 375-397. http://www.ascilite.org.au/ajet/ajet22/tsai.html

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research, 10*(2), 1-24.

Wong, R. (1993). Pronunciation Myths and Facts. *English Teaching Forum, 31*(4), 45-46.

Zheng, T. (2002). "Tell me more Chinese." Retrieved Month day, year, from the World Wide Web http://www.calico.org/CALICO_Review/ review/tmm-chinese00.htm

APPENDIX

Questionnaire on MyET

1. What features of this software program are helpful for improving your pronunciation?
   (Please fill in a number to indicate the extent of helpfulness:
   5 most helpful, 4 very helpful, 3 helpful, 2 a little bit helpful, 1 not helpful at all.)
   ___ (a) Overall score
   ___ (b) Correction of segment:
   ___display of phonetic symbols and scores
   ___ tips for pronunciation of segments
   ___ 3D animation of mouth
   ___ contrast display of spectrogram with model utterance
   ___comments
   ___ (c) pitch:
   ___ display of syllable and key
   ___ contrast display of pitch with model utterance
   ___ comments
   ___ (d) timing:
   ___ display of syllable and speed
   ___ comments
   ___ (e) emphasis:
   ___ display of emphasis
   ___ comments
   ___ (f) Recording design
   ___ (g) Clicking on the phonetic symbols or certain regions of spectrograms to listen to target sounds

2. By studying the acoustic spectrum are you able to understand the weak points in your pronunciation?
   Yes, I can. ___ No, I can't. ___
3. After practicing several times, are you able to be aware of the correctness (or incorrectness) of your own pronunciation?

   Yes, I can. ___ No, I can't. ___

4. Given more time to practice, do you think this program could help you improve your pronunciation?

   Yes, I think it could. ___ No, I don't think so. ___

5. Would you comment on or provide an overall evaluation of this software program?

   Yes, I would. ___ No, I wouldn’t ___