國立澎湖科技大學應用外語系

英文實務專題

Using Speech Recognition Systems to Improve Performance in Foreign Language Learning

運用語音辨識系統提升外語學習成效

學生姓名: 吳奕辰　1105402003

林琇敏　1105402008

李映宣　1105402010

徐靖怡　1105402011

指導老師: 譚峻濱　老師

中華民國一０八年十二月六日

Abstract

Because of the many changes witnessed during this era of explosive technological innovation, in which information is available readily on the Internet, more and more people are using 3C products. Automatic speech recognition (ASR) systems are applied in many kinds of 3C products, and people are using ASR software to learn foreign languages. Many people in the e-generation have attempted to improve their language proficiency with mobile or computer-assisted learning software. The purpose of this study was to apply an ASR to foreign language learning. It may be possible to use this software to allow students to practice their pronunciation autonomously and improve their oral expression skills. We selected a total of 26 students in the Department of Applied Foreign Languages at the University of Science and Technology to serve as our subjects and used the automatic speech analysis system (ASAS) referred to as MyET to perform a pre- and post-test comparison. The findings are as follows. The use of ASAS in English speaking did not improve learning outcomes. However, the ASR system did help learners improve their ability to speak English in computer language learning.

***Keywords*: E-learning, Mobile learning, ASAS (automatic speech analysis system), English speaking ability, corrective feedback**

1. Introduction

1.1. Background

In Taiwan's educational environment, the education policy for students who major in English emphasizes reading and writing rather than speaking skills. Therefore, students' speaking ability is relatively weak, despite the fact that many have studied English for many years, whether in school or cram school. Although they have basic reading and writing skills, they are afraid to speak English when they meet foreigners or answer questions in English in the class, even if only to talk to their foreign language teacher, and having a conversation with them is out of the question for them. Pronunciation is important, because it serves as a person's first impression. When a person wants to make friends with foreigners, and can speak their language fluently and accurately, it creates trust from the beginning, and they will be willing to communicate with the person, as they will believe that they are talking to a native English speaker.

There are many difficulties in learning pronunciation traditionally. In Taiwan's English education environment, some English teachers have a slight Taiwanese accent

when they speak English. This makes it difficult for students to learn the native American or British accent in school. Thus, with a Speech Recognition system and English Corrective Feedback Assistance System, students can learn pronunciation deeply not only in class, but also in their homes. Zhou (2005) proposed that teachers' accents may prevent them from teaching students correct pronunciation and simply studying online with social media will not improve their conversational skills and pronunciation. To solve such problems, a Computer Assisted Pronunciation Training System that uses Automatic Speech Recognition Technology (ASR) not only can provide a stress-free learning environment, but also allow students to practice repeatedly. It also provides feedback and correction functions for individual students' pronunciation problems.

This study used an automatic speech analysis system (ASAS), a mobile phone software that provides the user with immediate feedback to correct his/her speech. ASAS assesses the user's rhythm, intonation, and volume while speaking English, and allows repeated practice with practical sentences used in real life dialogue. In addition to the acoustic fingerprinting function, it can correct the pronunciation of a single character accurately to reach the correct emphasis in English. Using multimedia to construct an interactive and more diverse learning approach may involve listening and reading skills (Warsehauer, 1996). ASAS not only increases the user's convenience, but also makes it more fun to learn a foreign language.

1.2. Significance of the study

This study explored the way to use ASAS to improve students' performance in foreign language learning. The study's goal was to provide a useful tool for students and teachers in English education. This research can increase students' interest in learning English with ASAS, and improve their English proficiency by providing immediate corrective feedback, and teachers can use ASAS as a teaching aid to alleviate students' pronunciation problems. This type of research is conducted rarely. Therefore, our study outcomes can help scholars who wish to use ASAS to provide students with instant interactive learning opportunities, corrective feedback, and a stress-free learning environment.

1.3. Purpose of the study

The purpose of this study was to investigate ASR's effects on the ability of students who major in English to improve their pronunciation performance. Specifically, we investigated whether using ASR can improve students' pronunciation and increase their reading speed, learning efficiency, and autonomy.

1.4. Research questions

Research Question 1: What is the effect of using the English Corrective Feedback Assistance System on college students' English listening and speaking skills?

Research Question 2: What effect does using the English Corrective Feedback Assistance System have on college students' degree of immersion?

2. Literature Review

2.1 Communicative language teaching

Situational Language Teaching (SLT) was taught in the UK initially, but in the late 1960s, when the US began to review the theoretical basis of the Audiolingual Method (ALM), the UK began to perceive that communicative competence was the focus of language teaching. The ability to communicate echoes the linguistic competence the American linguist Chomsky described. The UK and the US share the same views and ideas: the communication and ability required to learn a language should be the syllabus for language teaching. Therefore, the British, American, and European parties developed Communicative Language Teaching (CLT), which has the following three learning principles: (1) communication; (2) task, (3) and meaning. The CLT features Nunan (1991) defined are recognized by the majority and have the following characteristics:

1. Communicate through language to achieve learning through interaction
2. Use life encounters during the learning process
3. Let students concentrate on language opportunities and learning processes
4. Integrate the student's personal experiences in the classroom
5. Try to combine the language outside the classroom with the context of the classroom learning content

As can be seen from the above, the CLT method's advantage is that students can express what they wish in the classroom. The method is effective in the student learning process and also can help students integrate their personalities directly into second language learning. In the teaching process, it usually is unnecessary for native speakers to practice pronunciation and language structure. Students are encouraged to learn from scratch and will not give up easily. In addition, because the teaching materials are related to society and life, students can learn the culture of the second language and various aspects of life.

Many studies today use communication pedagogy. Long and Richards (excerpted from Warschauer & Kern, 2000: ix) once said: “Because computer technology is between learners and teachers, and between second language users, communication has opened up new opportunities, and many language teachers have realized the potential of computer-aided language teaching.” Because of advances in technology, English learners now have more opportunities to learn from computer-assisted instruction on the Internet.

2.2 Game-based learning, GBL

Game-based learning (GBL) refers to learning through an Internet game platform, which allows students to solve problems in the game, overcome challenges, and achieve a sense of accomplishment. The learning style should take into account both gameplay and education to achieve both educational and entertainment goals.

In 1996, Hogle proposed that a game has the following advantages for learning: 1) It increases the motivation to learn and enhances the interest in learning. Because of curiosity and expectation, control and interaction, and the fascination of the game’s storyline, the interest and intrinsic motivation to learn can be improved. The player will feel a sense of accomplishment and will be willing to try again when faced with difficult challenges; 2) Compared to traditional courses, games have a relatively better effect on memory retention; 3) Many game learning software programs provide opportunities for practice, and allow the learner to repeat the operation and receive feedback in real time, so that the learner can self-assess his/her learning; 4) They promote high-level thinking by integrating teaching content into the game, so that the learner can solve problems, think, and make decisions in the game repeatedly. Learners should be able to find the solution by combining what they have learned. The teaching content will continue to enter their minds and is the best way to learn.

With the development of technology, Prensku (2007) also proposed the term Digital Game-Based Learning, which refers to using computer games to learn and acquire knowledge. Sandberg et al. (2011) designed game-based learning on smart phones to allow students to learn through various types of games. As more and more educators and researchers use game-based learning, Zhu (2012) introduced this method to an English course, and found that it increased students’ enjoyment of the course. Chang, Liang, Chou, and Li (2017) studied whether using game-based learning would make a difference in immersion

After a long period of research, the results to date have shown that game-based learning has no adverse effects on English learning, and enhances students’

motivation and interest in learning, and increases their immersion.

2.3 Corrective feedback

Corrective feedback refers to the response that the system provides when learning a foreign language. Corrective feedback is used widely in oral language learning, and is incorporated into the speech recognition system. Students can reduce their errors and make progress through self-learning without the need for teachers. It provides learners with an opportunity to discover their errors and make adjustments themselves. At present, most of the research on corrective feedback is based on Lyster and Ranta's study (1997), which summarized six types of corrective feedback (see Table 1).

Table 2-1 Corrective Feedback

Corrective feedback category	Definition	Example
Explicit Correction	The teacher points out the studentMyET's mistakes clearly and gives the correct answer to the mistake.	Student: I was at China two weeks ago. System: Not at China, in China. We say: "I was in China two weeks ago."
Clarification Request	Because of the way questions are asked, the statement is inaccurate or misunderstood.	Student: Which airline you want to? System: Sorry? Could you say that again?
Recast	The teacher reorganizes all or part of the student's sentences and corrects the mistakes.	Student: I was at Taiwan last month. System: In Taiwan.
Metalinguistic Clues	When a student makes a formal mistake, the teacher points out the wrong part according to the words the student spoke.	Student: I like to movie. System: Movie is a noun.

Elicitation	When the student makes a mistake, the teacher guides him/her directly to speak the statement correctly.	Student: It's hot in winter. System: It's hot in…? Student: Summer.
Repetition	The teacher repeats the learner's incorrect sentence to attract his/her attention.	Student: My favorite teacher was Suzuki. System: My favorite teacher WAS Suzuki?

2.4 Speech recognition

With the rapid developments in science and technology, people have made breakthroughs in the field of artificial intelligence and speech recognition, and introduced this technology into people's lives directly, for example, Apple's built-in artificial intelligence assistant software, Siri, Google's national translation program, and Amazon's smart assistant, Alexa. These intelligent assistance programs use voice recognition to make people's lives more convenient. Voice recognition systems can be used to listen to various data, set phone functions, direct phone communications, and translate. They even can tell the user jokes.

The traditional teaching method is a face-to-face method. All sources of knowledge derive from the textbooks in the classroom, and the students' breakthroughs are limited by the textbook's guidance. However, because digital learning courses are available readily now, ASR has witnessed great progress in the past decade. Many researchers and experts and scholars have expressed their views as follows:

Warschauer (1996) stated that speech recognition technology can be used to construct more interactive learning methods. Shadiev, Hwang, and Huang (2006) described the analysis of college students' perceptions and experience using continuous ASR. Neri, Mich, Gerosa, and Giuliani (2008) indicated that speech recognition is a better way for learners to improve their English pronunciation in a shorter period of time than is traditional instruction. Morton and Jack (2010) applied voice computer interaction into auxiliary language learning. Kuo, Shadiev, Hwang, and Chen (2012) demonstrated that supporting academic activities in an online synchronized learning environment can assist both English and non-English speaking learners effectively. Shadiev et al. (2015) used speech recognition technology to allow non-native English speakers to learn in an English seminar. Ahn and Lee (2016) designed a system used

to learn spoken English with speech recognition on mobile phones. Their results showed that learners believe that speech recognition helps them learn and even makes learning interesting.

In summary, the application of speech recognition systems to teaching is a subject worthy of study. There are many resources on the Internet, which can be regarded as a large textbook library with diverse content, which not only attracts learners' attention, but also enhances teachers' teaching quality to increase students' interest to learn actively.

3. METHODOLOGY

3.1. Introduction

This section describes the methodology adopted to use ASAS to improve students' performance in foreign language learning. First, our research included applied senior foreign language students as participants. Second, we used ASAS-MyET and a questionnaire as instruments. Third, our procedures included testing for four weeks and a questionnaire to explore the effects of learning on MyET. Fourth, our data analysis showed that ASAS could improve the students' pronunciation and accent.

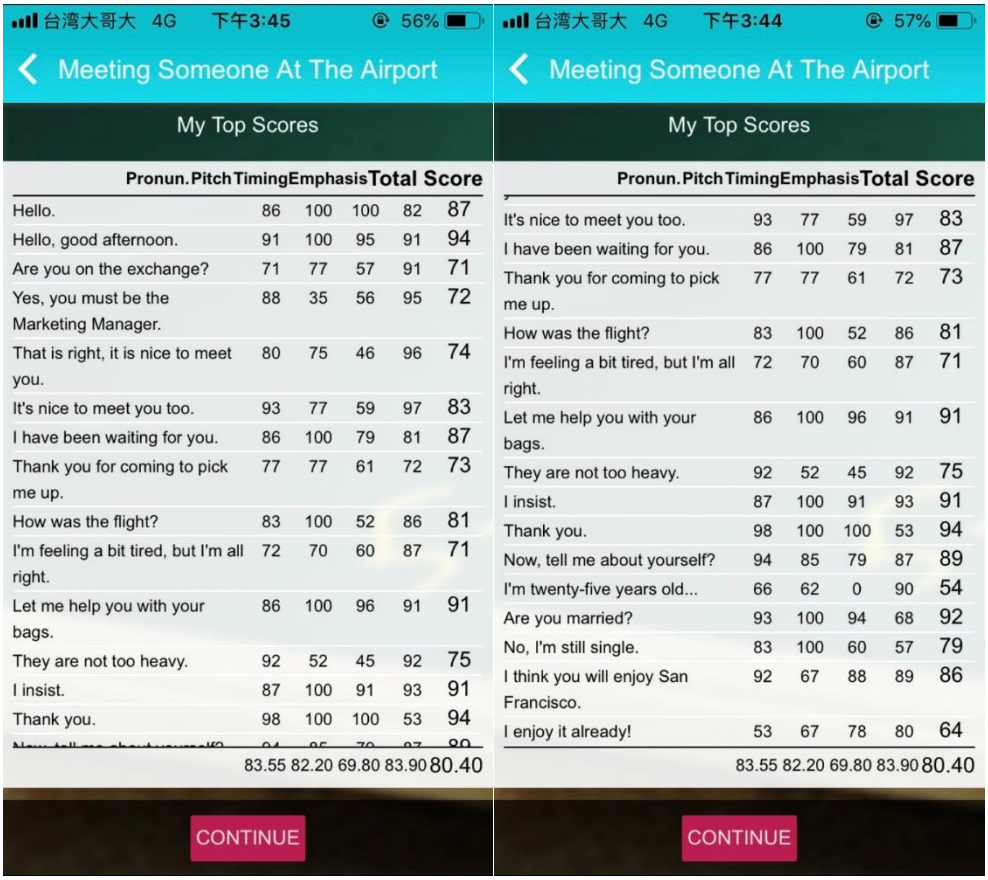
3.2. Participants

The study included twenty-six university students at National Penghu University of Science and Technology, eleven males and fifteen females. 20 students began to learn English during their first year of primary school and 6 began in kindergarten. 24 students had a TOEIC score over 750 (2), 550-750 (17), and 350-550 (5); 8 had not taken the TOEIC.

3.3. Instrument

We used the ASAS–MyET to investigate learning effectiveness. MyET uses the most advanced cloud architecture, so that students could use the same account and use MyET on Windows PC, an iPhone, iPad, Android Phone, Android Pad, and Mac OS, and their results were sent automatically to the cloud server so the teacher could check them at any time. The system included Aviation English, sightseeing English, entertainment English, English for international students, English for emergencies, and English for shopping (Figure 3-1). The closed-end questionnaire, the main purpose of which was to understand the students' immersion and feelings about this system, was used as a reference for future teaching.

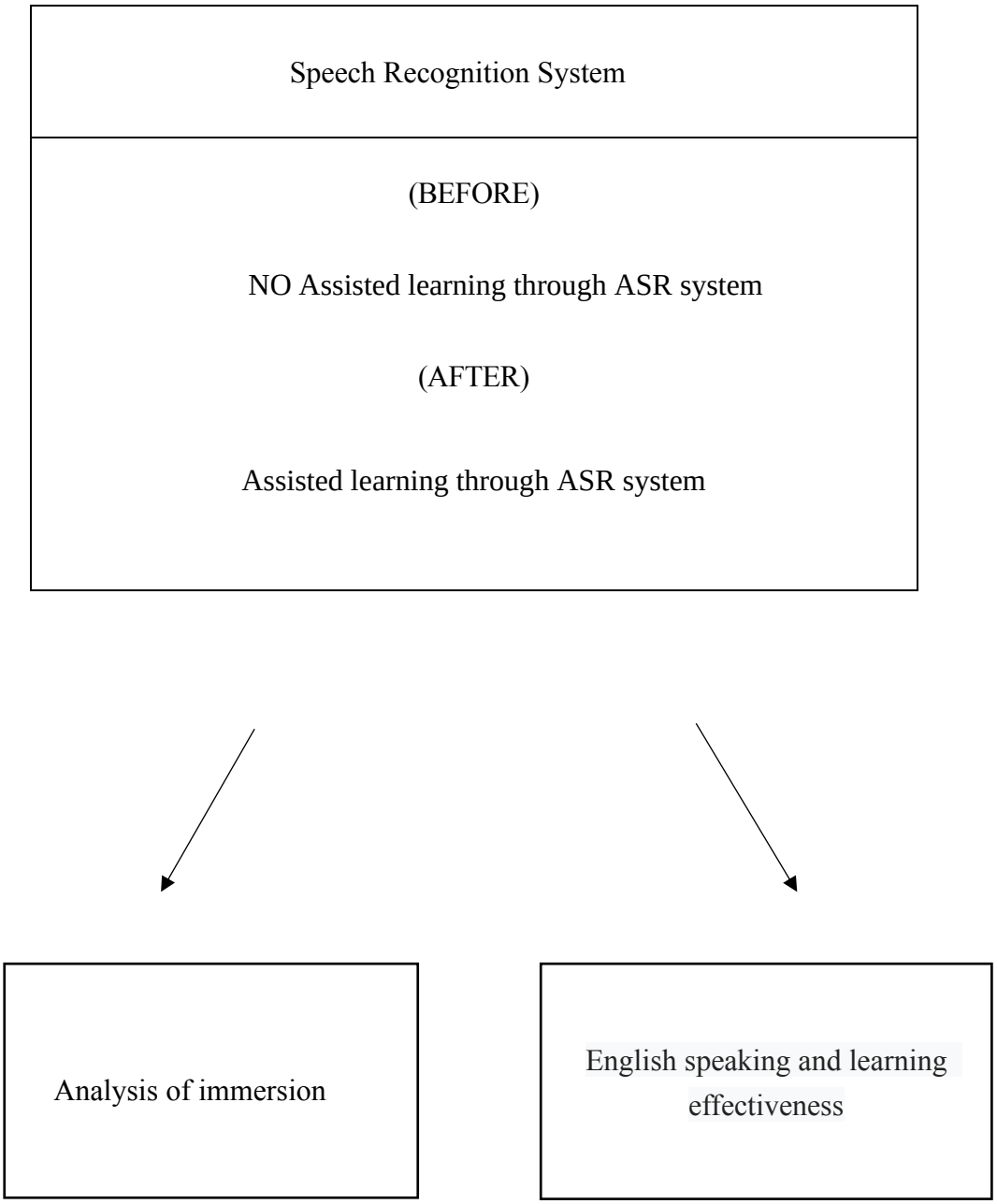
Table 3-1 The Lesson: Meeting Someone at the Airport in MyET



3.4. Research structure

This study focused on game-based teaching methods. Through corrective feedback, learners can learn to speak English through the system's voice recognition function, which allows them to understand pronunciation, fluency, and whether the intonation and volume are correct; the process also uses the inability to respond to indicate that the statement was incorrect or misunderstood to allow the learner to make corrections, and gives feedback on the criteria above as the basis for the final grade, thereby promoting the learner's ability to speak English, and the effectiveness of their learning. Therefore, this study included a pre-test and post-test, with which we determined whether the students' English improved with the aid of a speech recognition system.

Figure 3-1



3.5. Procedures

We invited the seniors in our department to be our experimental subjects, and used the 4-week, fourth-grade elective course—Aviation English—to perform this experiment. After the aviation English course, the students could simulate an English aviation

conversation with MyET. The context of the conversation was checking-in at the airport counter. First, we provided 5 minutes of operating instructions, so that the subjects would become familiar with, and know the way to operate the system. The pre-test began 10 to 15 minutes later. The students used the system for 10 to 20 minutes from the second to the fourth weeks to learn, practice, and correct their speech. In the second week, they studied and practiced the aviation clearance situation, such as passing through immigration and customs, and practiced a conversation at the check-in counter at the airport during the third week. The 15-minute post-test was given in the fourth week, and the post-learning questionnaire was administered during 10 minutes after the post-test. The questionnaire included the students' immersion and learning experience with the system.

3.6. Data analysis

This study was based largely in the applied foreign language department at Penghu University of Science and Technology. We wanted to explore two different groups, and used SPSS to analyze the data and the questionnaire to investigate whether students demonstrated any differences in immersion.

4. Results and Discussion

4.1. Demographic variables

The 26 subjects in this study were students in the senior grade in the Department of Applied Foreign Languages at National Wuhu University of Science and Technology. They ranged in age from 22 to 23, and all were educated in English. 19% had a TOEIC score of 350-550, 65% a score of 550-750, and 8% a score over 750; 8% did not take the TOEIC test. In addition, 23% of the students began to learn English in kindergarten, and 77% began in the first year of primary school.

Table 4-1

TOEIC score				**Time subjects began to learn English**	
No participation	**350-550**	**550-750**	**Over 750**	**Kindergarten**	**First year of primary school**
8%	19%	65%	8%	23%	77%
N=2	N=5	N=17	N=2	N=6	N=20

4.2. Learning effectiveness analysis

This study explored the effectiveness of oral English learning, so the results of previous tests were the primary basis of comparison. The test scores in this lab include the "fluency" test, "tongue" test, "pronunciation" test and "volume" test. The total score for each is 100 points. The subjects' time, textbook, and scope were the same. The differences between pre- and post-test were analyzed according to the scores on each test, and whether the use of MyET software improved the students' oral English performance.

Section one: The post test code reliability

The significance test is a method to determine whether there is a difference between the effects of an experimental treatment group and control group or two different treatments, and whether the difference is significant. If the p value is <0.05, there is a significant difference between the two groups.

Section two: The ability to speak and learn

Because English listening and speaking affect each other in the English learning process, this was included in the test to understand the students' comprehensive ability in listening to, and speaking, English. Therefore, the independent sample t-test was performed to analyze the difference between the test subjects' pretest and post-test scores on English performance. Therefore, the scores on the "fluency," "tongue," "pronunciation," and "volume" tests were analyzed with an independent sample t-test. The results of the fluency analysis (Table 4-2) showed that the subjects' mean pre-test score was 76.27 (SD = 6.26), while their mean post-test score was 78.48 (SD = 6.18), because of the two groups' variation in Levene's test. There was a significant difference in the homogeneity test ($F = 0.10$, $p = 0.76$), so the variance was not equal. The value of the independent sample t-test was -1.28 ($p = 0.21$). There was no significant difference between the speech recognition group and the speech recognition corrective feedback group. Therefore, it can be inferred that neither group of subjects made progress in English fluency after the experiment.

Table 4-2 Fluency Independent Samples *t*-verification

Construct	Group	N	M	*SD*	*t*	*p*
Fluency	Before	26	76.27	6.26	-1.28	0.21
	After	26	78.48	6.18		

The results of the intonation analysis (Table 4-3) showed that the mean pre-test score was 85.21 (*SD* = 3.29), and the mean post-test score was 84.90 (*SD* =3.94) because of the variation between the two groups in Levene's test. There was a significant difference in the homogeneity test (F = 1.42, *p* = 0.24), so the variance was not equal. The *t*-value of the independent sample was 0.31 (*p* = 0.76). There was no significant difference between the speech recognition group and the speech recognition corrective feedback group. Therefore, it can be inferred that the two groups of subjects made no progress in English intonation before and after the experiment.

Table 4-3 Intonation Independent Samples *t*-verification

Construct	Group	N	M	*SD*	*t*	*p*
Intonation	Before	26	85.21	3.29	0.31	0.76
	After	26	84.90	3.94		

The results of the pronunciation test (Table 4-4) showed that the mean pre-test score was 83.33 (*SD* = 3.80) and the mean post-test score was 84.79 (*SD* = 4.44), because of the difference between the two groups in Levene's test. There was a significant difference in the homogeneity test (F = 0.37, *p* = 0.55), so the variance was not equal. The *t*-value of the independent sample *t*-test was -1.27 (*p* = 0.21). There was no significant difference between the speech recognition group and the speech recognition corrective feedback group. Therefore, it can be inferred that neither group of subjects made progress in English pronunciation before and after the experiment.

Table 4-4 Pronunciation Independent Samples *t*-verification

Construct	Group	N	M	*SD*	*t*	*p*
Pronunciation	Before	26	83.33	3.80	-1.27	0.21
	After	26	84.79	4.44		

The results of the analysis of volume (Table 4-5) showed that the mean pre-test score was 84.23 (*SD* = 2.62), while the mean post-test score was 83.15 (*SD* = 3.21) because of the variation in the two groups in Levene's test. There was a significant difference in the homogeneity test (F = 0.79, *p* = 0.38), so the variance was not equal. The *t*-value of the independent sample *t*-test was 1.33 (*p* = 0.19). Hence, there was no significant difference between the speech recognition group and the speech recognition corrective feedback group, and it can be inferred that the two groups of subjects made no progress in English volume before and after the experiment.

Table 4-5 Volume Independent Samples *t*-verification

Construct	Group	N	M	*SD*	*t*	*p*
Volume	Before	26	84.23	2.62	1.33	0.19
	After	26	83.15	3.21		

4.3. Analysis of immersion

This study also explored whether the subjects' pre- and post-test scores on speech recognition exhibited differences in immersion during the learning process and satisfaction overall. Therefore, the questionnaire Wang and Hsu (2014) developed was used to collect the data. The responses to the questionnaire were scored on a five-point Likert scale that ranged from 1, *Agree strongly* to 5, *Disagree strongly*. Because the study was designed to investigate the effect on the immersion level overall, in the immersion section, the analysis was performed on 6 facets using independent sample *t*-tests.

Enjoyment

The results showed that the *t* value was 0.72 (p = 0.48), indicating that Enjoyment did not differ significantly between the pre-test and post-test (Table 4-6).

Table 4-6 Enjoyment

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Enjoyment	Before	0.93	13	2.38	0.97	0.72	0.48
	After		13	2.13	0.84		

Telepresence

The results showed that the *t*-value was 0.96 (p = 0.35), indicating that there was no significant difference in Telepresence between the pre- and post-test (Table 4-7).

Table 4-7 Telepresence

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Telepresence	Before	0.83	13	2.21	0.79	0.96	0.35
	After		13	1.90	0.85		

Focused Attention

The results showed that the *t*-value was 0.44 ($p = 0.66$), indicating that there was no significant difference in focused attention between the pre- and post-test (Table 4-8).

Table 4-8 Focused Attention

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Focused attention	Before	0.90	13	2	0.78	0.44	0.66
	after		13	1.87	0.70		

Involvement

The results showed that the *t*-value was 1.34 ($p = 0.19$), indicating that there was no significant difference in the degree of involvement between the pre- and post-test (Table 4-9).

Table 4-9 Involvement

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Involvement	Before	0.97	13	2.10	0.76	1.34	0.19
	After		13	1.72	0.71		

Time Distortion

The results showed that the *t*-value was 1.59 ($p = 0.12$), indicating that there was no significant difference in time distortion between the pre- and post-test (Table 4-10).

Table 4-10 Time Distortion

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Time distortion	Before	0.84	13	2.48	0.84	1.59	0.12
	After		13	1.98	0.76		

The five facets were analyzed further to determine whether there was a significant difference between any two. Therefore, the independent sample *t*-test was used for analysis. The analysis showed that, regardless of the facet in the degree of immersion, there were no significant differences between the pre- and the post-test.

Satisfaction

Finally, the independent sample t-test was used to analyze students' satisfaction. The *t*-value was 1.62 ($p = 0.12$). Thus, there was no signification difference in satisfaction between the pre- and post-test, as shown in Table 4-11.

Table 4-11 Satisfaction

Construct	Group	Cronbach's alpha	N	M	*SD*	*t*	*p*
Satisfaction	Before	0.91	13	2.19	0.82	1.62	0.12
	After		13	1.77	0.47		

4.4. Comprehensive Discussion

This study proposed to use a ASR software system that included corrective feedback to promote the ability to learn to speak English, and used the game context as a background, so that learners would be interested to explore the system to improve their ability to speak in English. We also explored whether the ability to learn can promote immersion. Therefore, this section will delve into the effects and outcomes of the subjects' pre- and post-tests on the effectiveness of immersion and English speaking ability.

First, we discuss the effect of the learning system combined with speech recognition and corrective feedback on learning outcomes. The pre-test in this study allowed the students to practice learning activities before they began them. After the experimental subjects conducted the learning activities in English speaking ability, the independent sample *t*-test was used to analyze the post-test results to understand the students' speech in the English tongue, or the difference in learning effectiveness.

In the post-test results section, the statistical analysis showed that there were no significant differences between the subjects in the "fluency," "tongue," "pronunciation," and the "volume" tests. The speech recognition corrective feedback

used in MyET software gives the student corresponding corrective feedback when errors occur and allows the subject to use the clarification request in the learning process, correct his/her sentence, and gives the learner four items. The test score allows the subject to correct the errors against the ratings the system provides. Therefore, it can be inferred that speech recognition combined with corrective feedback helps students think about and correct their errors, and enhances their English learning ability. However, because the subjects in this study needed to perform the post-test experiments in a short period of time, the results were not significant.

The study also allowed subjects to comment on their learning using the speech recognition corrective feedback system, two of whom said:

Student A: "I can practice my speaking ability."

Student B: "It's fun, but it will be better to have [a] longer time."

Second, we discuss the effect of the speech recognition combined with corrective feedback learning system on immersion and satisfaction. According to the statistical results in Section 3, there were no significant differences in the subjects' degree of immersion and satisfaction.

With respect to immersion, we divided the content into six facets and analyzed them using independent sample *t*-tests. Because the learning activities in this experiment were all based on the learning system, regardless of whether the subject was in the pre- or post-test phase, they all were learning in the same game theme. The only difference was that the subject was only allowed to use the system. The content of the grading is practiced and improved, so the differences in this mechanism may have an effect on learning outcomes. Although the facets in the degree of immersion and satisfaction were not statistically significant, from the learners' feedback on the use of the speech recognition corrective feedback system learning, we can see that some learners believed the system enhanced their ability to speak. However, because of the influence of time and environment, there will be errors in the experimental results, such as speech recognition in different environments, poor telephone reception, etc., or the length of time between the pre- and post-test that may affect the experimental outcomes. Therefore, environmental factors affected the accuracy of identification, and the time was not mastered well, and thus affected the students' immersion and satisfaction. We also learned from Blanka Klimova's (2019) research that the application of the ASR system to a mobile phone may require teachers' guidance to enhance learning outcomes. Further, Kasrani, Vasoya, Shivakumar, and Pei's (2018) study found that ASR combined with corrective feedback on the computer is effective in improving foreign language learning ability. Therefore, computer language learning

may be more effective than learning on a mobile phone.

5. Conclusion

This study adopted game-style learning as the context, and used a pre-test/post-test design without (pre-test) and with (post-test) a tutoring and speech recognition system to understand whether the auxiliary teaching using the speech recognition system can improve students' performance significantly. The paper summarizes and describes whether this study improved students' performance in learning English with a ASR analysis system.

In this study, through the speech recognition system, students who are afraid to speak English were able to use game-style teaching to achieve learning in an easy and interesting way.

The study measured the subjects' fluency, intonation, pronunciation, and volume during the learning process. When the subject makes an error, the system detects it through speech recognition, and finally allows the subject to understand their learning situation. We used the pre- and post-test results to explore the system's effectiveness in learners' performance and their immersion in learning English. The study used a quasi-experimental method to collect data and analyzed it through statistical methods. The results were as follows:

1. What is the English speech recognition correction feedback system's effect on college students' ability to speak English?

During the study, the subjects corrected their pronunciation, fluency, intonation, and volume through the speech recognition system. After using the independent sample *t*-test, the results showed that the students' ability to speak English using the speech recognition corrective feedback system did not improve. The study also showed that the subjects' immersion affected their performance in the learning process.

2. How does the use of an English speech recognition correction feedback system affect college students' immersion?

According to the findings above, there was no significant difference in the students' degree of immersion, and thus, the students made no progress in fluency, pronunciation, intonation, and volume. The study found that the subjects may not have been invested fully in the speech recognition system because of the time distortion, which led to no significant difference in the pre- and post-test results.

References

Ahn, T. Y., & Lee, S. M. (2016). User experience of a mobile speaking application with automatic speech recognition for EFL learning. British Journal of Educational Technology, 47(4), 778-786. doi:10.1111/bjet.12354

Blanka Klimova. (2019). Impact of Mobile Learning on Students'Achievement Results DOI:10.3390/educsci9020090

Chang, C.-C., Liang, C., Chou, P.-N., & Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-gamebased learning? Perspective from multimedia and media richness. Computers in Human Behavior, 71, 218-227. doi:10.1016/j.chb.2017.01.031

Cucchiarini, C., Neri, A., & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. Speech Communication, 51(10), 853-863. doi:10.1016/j.specom.2009.03.003

Ellis, R. (2001). Introduction: Investigating Form-Focused Instruction. Language Learning, 51(s1).

Faqeih, H. I. (2015). Learners' Attitudes towards Corrective Feedback. Procedia - Social and Behavioral Sciences, 192, 664-671. doi:10.1016/j.sbspro.2015.06.101

Kuo, T. C. T., Shadiev, R., Hwang, W. Y., & Chen, N. S. (2012). Effec of applying STR for group learning activities on learning performance in a synchronous cyber classroom. Computers & Education, 58(1), 600-608. ts doi:10.1016/j.compedu.2011.07.018

Kasrani, Imen M.; Vasoya, Miteshkumar M.; Shivakumar, Ashutosh; Pei, Yong.((14th, Lisbon, Portugal, Apr 14-16, 2018) A Mobile Cloud Computing Based Independent Language Learning System with Automatic Intelligibility Assessment and Instant Feedback

Leeman, J. (2003). Recasts and Second Language Development. Studies in Second Language Acquisition, 25(01). doi:10.1017/s0272263103000020

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. Studies in Second Language Acquisition, 37-66.

McDonough, K. (2005). Identifying the Impact of Negative Feedback and Learners' Responses on Esl Question Development. Studies in Second Language Acquisition, 27(01). doi:10.1017/s0272263105050047

Morton, H., & Jack, M. (2010). Speech interactive computer-assisted language learning: a cross-cultural evaluation. Computer Assisted Language Learning, 23(4), 295- 319. doi:10.1080/09588221.2010.493524

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children.

Computer Assisted Language Learning, 21(5), 393-408. doi:10.1080/09588220802447651

Prensky, M. (2007). Digital Game-Based Learning. New York: McGraw-Hill.

2.HogleJ. G. , （1996）「Considering Games as CognitiveTools: In Search of Effective "Edutainment" University of Georgia Department of Instructional Technology,」

Rahimi, M., & Zhang, L. J. (2015). Exploring non-native English-speaking teachers' cognitions about corrective feedback in teaching English oral communication. System, 55, 111-122. doi:10.1016/j.system.2015.09.006

Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: An evidencebased study with fifth graders. Computers & Education, 57(1), 1334-1347. doi:10.1016/j.compedu.2011.01.015

Shadiev, R., Hwang, W.-Y., & Huang, Y.-M. (2006). Liberated learning: Analysis of university students' perceptions and experiences with continuous automated speech recognition. Journal of Instructional Science and Technology, 9, 1-19.

Shadiev, R., Hwang, W.-Y., Huang, Y.-M., & Liu, C.-J. (2015). Investigating applications of speech-to-text recognition technology for a face-to-face seminar to assist learning of non-native English-speaking participants. Technology, Pedagogy and Education, 25(1), 119-134. doi:10.1080/1475939x.2014.98874

Shadiev, R., & Huang, Y.-M. (2016). Facilitating cross-cultural understanding with learning activities supported by speech-to-text recognition and computer-aided translation. Computers & Education, 98, 130-141. doi:10.1016/j.compedu.2016.03.013

Warschauer, M. (1996). Computer Assisted Language Learning: an Introduction. Multimedia language teaching, 3-20.

Warschauer, M., & Kern, R. (Eds.). (2000). Network-based language teaching: Concepts and practice. Cambridge, England: Cambridge University Press.

Zhou（2005）. The application of speech recognition in correcting pronunciation