Speech-to-text recognition in University English as a Foreign Language Learning

Kate Tzu Ching Chen[1]

Abstract

This study explored the potential of adopting speech-to-text recognition (STR) technology for English as a foreign language (EFL) oral training in class at the university level. An action study method was set to investigate the effects of implementing STR app tasks for EFL oral training with 27 students in one class for one semester at a university located in Taiwan. Data were obtained through pre- and post-tests, speaking practice results, field-observation notes, student reflective journals, and end of class survey. The results of the quantitative data analysis indicated that the STR app tasks were effective in increasing students' English speaking ability. Students also expressed positive attitudes toward the use of the tasks in the STR app. Further, the qualitative data analysis showed that the students found these tasks highly motivating and can quickly engage them in learning. The STR app tasks that provide repetitive training of spoken English, particularly on pronunciation, fluency, and vocabulary acquisition approved be more beneficial to learners than traditional teaching methods. It is highly recommended for teachers to design a variety of STR tasks to meet individual learners' needs and preferences.

1 Introduction

Good English speaking ability is very important for non-native English speakers to communicate well and efficiently with people around the world; however, this skill is difficult to acquire for most Taiwanese university students (Chen, 2011). Many

---

[1]   Department of Applied English, Chaoyang University of Technology, 168 Jifong E. Rd., Wufong District, 41349 Taichung City, Taiwan (R.O.C.)

students lack the ability to speak and express themselves in fluent English because English education in Taiwan pays less attention to oral training. Students also generally lack the confidence to speak in English because of the influence of the Chinese culture. When asked to speak in English in class, students often experience excessive anxiety and fear of being insulted or laughed at by their peers (Kung, 2017).

Furthermore, the influence of the local dialect makes it more challenging to learn the intricacies of speaking English natively (Pal, 2013). The differences between local dialect and target language, in this case, Chinese and English are also the reasons why English oral skills are difficult to master. For example, the intonation specifies emotion in English, but meaning in Chinese. Stress is another example as Taiwanese EFL learners often misplace the stress in English words and sentences which cause the interruption and unintelligibility in the conversations (Bian, 2013). In fact, in typical Taiwanese oral English training courses, lecturers often find themselves limited by traditional teaching methods and materials, which are now inadequate for the current digital generation (Hsu, 2013). Students oftentimes lose their interest and focus in class and suffer from poor learning due to ineffective teaching methods. This is the reason why one area that needs more focus is technological tools to help improve the students' learning outcome and experiences.

Mobile devices have been widely recognized as effective and efficient tools for extending students' language learning (Heil et al., 2016). The use of mobile devices for language learning creates a learning environment conducive for interaction and collaboration, provides a variety of ways for learning, and allows quick access to all types of information; all of which could enhance learning outcomes (Wankel & Blessinger, 2013). It also enables students to engage with more effective learning experiences in and out of the class, allowing them to become dynamic, innovative, and knowledgeable in this changing digital era. Recently, the widespread penetration of mobile devices also led to the rapid development of language learning applications (Rosell-Aguilar, 2017). The availability of these apps offered a variety of educational activities that are thought-provoking, motivating, relevant, and meaningful, giving learners more opportunities to practice and enhance their language skills (Heil et al., 2016). Considering the popularity and convenience of mobile devices. Instead of preventing students from using and being distracted in class by mobile phones, teachers can now create class activities using this effective educational tool in their English classes (Chen & Hsu, 2020). Therefore, the idea of using mobile apps to assist students' oral training arose. Especially that researchers have urged that more studies are still needed to fully investigate the overall effectiveness of language learning apps (Rosell-Aguilar, 2017).

When speech-to-text recognition (STR) technology is used on mobile devices, it has been proven to help students to learn English better by increasing their learning performance, independent learning opportunity, and self-confidence (Shadiev et al., 2014). Yet, most of the existing research has focused mainly on using STR technology to assist in understanding English lectures, improving listening comprehension when talking to native English speakers, or enhancing English reading comprehension (Shadiev et al. 2017a, b). It has been shown that using STR technology has great potential as an English learning tool in class to improve EFL students' speaking proficiency. For example, Ahn & Lee (2016) found that students have expressed

positive attitudes toward the use of STR technology for speaking practices and were particularly interested in its unique function of immediate demonstration of learner's speech inputs. However, few studies, if any, measured the improvements by using STR for English speaking training which indicates that this application has not yet been properly explored and tested.

The use of mobile learning has been continuously proposed by scholars over the years. Some have even believed that educators, particularly those in higher education, should keep up with the world through mobile devices and social networks, and use technology as a bridge to connect formal and informal learning (Kukulska-Hulme, 2012). Especially considering that much of the teaching in oral training classes are repetitive and routine that causes students to lose their interest in learning. It is hoped that students can take the advantage of mobile technology for EFL oral training in class to make serious and boring courses more interesting, reduce students' learning pressure, and help achieve the learning outcomes without sacrificing the learning objectives. The goal is to introduce STR technology on mobile devices to encourage uses in class as a learning tool, and not regard it as a distraction. Therefore, this research utilized mobile apps with speech recognition technology to create learning tasks that could generate immediate learning results to help EFL university students improve their English oral skills. Through this, students will be encouraged to use their mobile devices for self-learning even after their class and continue to be lifelong learners. In this study, the following research questions were explored:

(1) How effective are STR app tasks in enhancing students' English oral proficiency?

(2) How do students perceive the STR app as an English Oral Training tool?

2 Literature Review

2.1 Mobile assisted language learning for EFL

Many researchers have suggested that a mobile device can be an ideal tool for language learning because of its availability, portability, and flexibility Jamaldeen et al., 2018; Niño, 2015; Persson & Nouri, 2018; Rosell-Aguilar, 2017; Wankel & Blessinger, 2013). Especially after 2007, when Apple launched their groundbreaking technology product, iPhone, which provided a variety of functions that are beneficial in enhancing learners' language skills (Godwin-Jones, 2011). Since then, many scholars have devoted their studies to the application of mobile assisted language learning (MALL), hoping to use this innovative technology to improve students' vocabulary learning (Chen & Chung, 2008; Huang et al., 2012), reading comprehension (Hsu et al., 2013), speaking proficiency (Gromik, 2012), and pronunciation and writing skills (Jia et al., 2012).

Among the applications of MALL, mobile apps have become the most popular and valued tools to be incorporated into language instruction in the past decade (Godwin-Jones, 2011). The mobile apps have both entertaining and educational functions which can especially arouse learners' learning motivation. It no longer merely offers cognitive and receptive language skill practices, but provides a full language learning

experience (Rosell-Aguilar, 2017). Therefore, it has been used to improve traditional in and out of classroom teaching and learning by increasing students' willingness to learn actively to achieve better learning results. In addition, since individual learners can decide when and how to use mobile apps for language learning, it is ideal for informal learning and to encourage and develop learning autonomy (Godwin-Jones, 2011).

In a rapidly changing society, it is especially necessary for higher education to integrate the widely used MALL into formal education to better engage students in active learning (Kukulska-Hulme, 2012). Several studies have demonstrated the effectiveness and benefits of MALL in the context of EFL in Taiwan. For example, Chen and Lin (2018) attempted to integrate learning tasks using mobile technology into EFL classes and found that Taiwanese college students showed positive attitude toward their language skills gains, especially on vocabulary and writing. A study that utilized mobile game-based learning to improve EFL vocabulary learning performance by Chen et al., (2019) revealed that game-related functions were positively correlated with vocabulary learning performance. Students mostly agreed that it was effective and satisfying for them in learning English vocabulary. However, despite the many benefits of MALL, it has not yet been able to be successfully applied in Taiwanese formal educational system. One of the key factors of this issue is that teachers are reluctant and incapable of incorporating MALL into classroom teaching (Hsu, 2015).

2.2 Speech recognition technology in English learning

The speech recognition technology is also known as automatic speech recognition, or computer speech recognition technology. Its goal is to automatically convert spoken human speech into corresponding text, initially using computers, and now through mobile devices (Hansen & Bořil, 2018). Many scholars have advocated that when learners use speech recognition as a learning tool, motivation and interest in the learning content are often enhanced. Learners show higher self-confidence which makes it easier for them to achieve the teaching goals of the educator, leading to lifelong learning (Handley, 2009).

In the field of STR technology for English learning, most scholars presented their study results on the effectiveness of speech recognition application for non-native English learners in supporting their lecture comprehension. Among them, Shadiev and his associates (2015) explored the use of STR technology to improve non-native English speaking students' understanding of the seminar content taught in English. The results showed that STR technology effectively helped students understand the seminar content, and that most students perceived the STR technology positively. Shadiev & Huang (2016) used the STR technology in the Windows operating system to improve the understanding of non-native English speakers of academic lectures. The research found that the spoken manuscript produced by the STR technology assisted the participants to have a more complete understanding of the lecture content. In another study, Shadiev et al., (2017a, b) proposed that STR technology can reduce the anxiety of non-native English speakers, thereby improving the effectiveness of language learning.

In the last decades, different STR applications have been investigated in English oral training with positive results. For example, Strik et al., (2009) combined the con-

cepts of computer-assisted language learning (CALL) and speech recognition, and developed a computer system to help language learners in oral proficiency training. Zhang & Liu (2018) also used computer speech recognition technology in the students' oral English training and concluded similar results. In addition, Coniam (1998) used Dragon Naturally Speaking as a tool for assessing one's English speaking ability, and the results showed that although speech recognition has yet to be developed, it proved that its accuracy is no less than that of traditional oral exams. Later in 2012, a study reported by Cox and Davies further confirmed that STR technology is well developed and can be used in oral English tests.

Currently, there is little credible research findings for using STR technology on mobile learning apps for improving the English speaking ability of EFL students. Among the few studies completed is the one by Ahn & Lee (2016) which developed an app to improve Korean students' English speaking proficiency. They concluded that mobile speaking application with automatic speech recognition has great potential for helping EFL students to learn to speak English. Despite of the potential use of STR technology in the academe, it is surprising that not many researchers have created classroom activities using this technology (Carrier, 2017). Most educators are eager to learn on how they could fully utilize new technology in the classroom setting to improve their own teaching methods and students' learning outcomes. In view of this, this research attempted to explore the effects of STR mobile learning apps on oral English training for EFL students.

3 Method

Ever since the teachers-as-researchers movement was established by the higher education institutions, action research has become a popular method for teachers to transform themselves as researchers to improve their teaching practices (Elliott, 1991). This study adopted this particular research method to explore the effects of implementing STR apps for oral English training and to improve this course for EFL students at the university level.

3.1 Participants

A total of 27 first year university students were recruited in an English oral training course at a technical university in Taiwan. There were 7 males and 20 females. The age ranged from 17 to 19 years old. All participants were intermediate-level English learners who have studied English for at least 9 years. None of them had previous experience of using STR apps.

3.2 The design and process of STR app tasks

The adoption of MALL and STR has proved beneficial for improved learning and that speech recognition technology has been widely utilized by mobile devices to enhance learning (Shadiev et al., 2014). Within the framework of MALL and STR, several mobile device apps that used speech to text recognition technology were

investigated. After reviewing numerous apps, two were decided on that were suitable for ESL learners and for them to develop and enhance their speaker-adaptive skills as well as pronunciation. These two apps were Speechy for iOS and SpeechNotes for Android.

Both apps are designed to transcribe spoken words into typed text by analyzing the speech inputs and using its algorithms to identify the best fit word. Without having any association with any app developers and/or its development, and through the researcher's own trial and exploration of these apps, found that some of its functionality is useful in teaching English speaking. In the process, participants were required to verbally read the assigned reading texts into their cell phone's microphone, then the app converts and transcribes the spoken phrases and sentences into typed words that provide immediate dictational results.
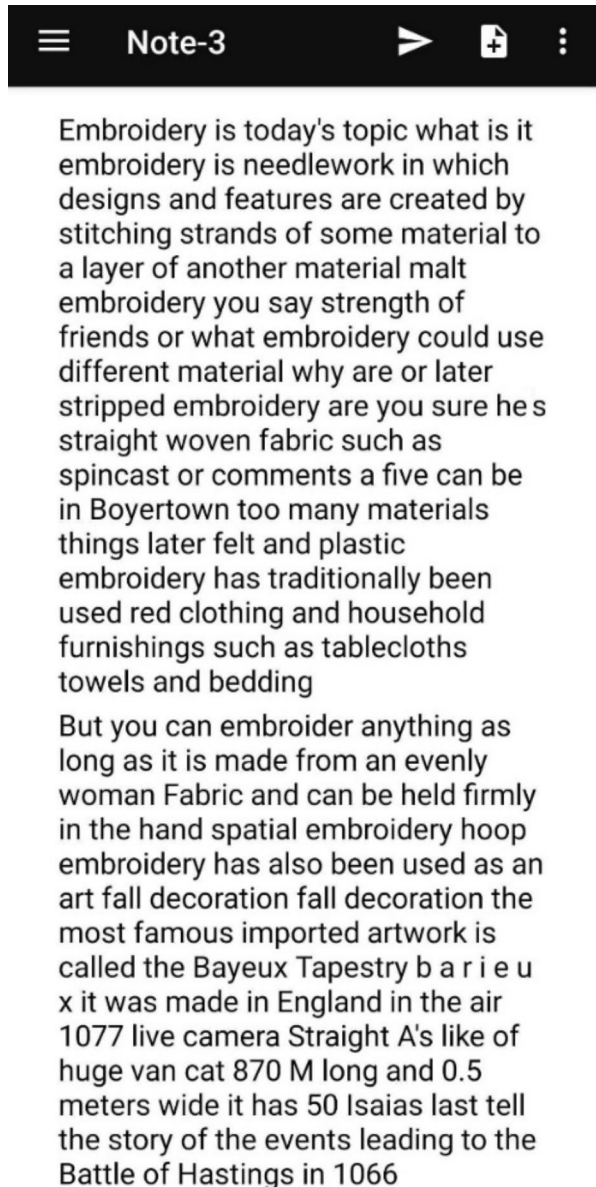
The learning tasks designed by the author were intended to utilize the functions of STR app within the MALL and STR framework for enhancing learners' speaking skills and this has become a promising research focus (Ya et al., 2013). The tasks were planned so that the use of the STR apps was a study requirement, generating consistent results, and found to be most suitable for students in their speaking practice class. The STR app was used to fulfill the students' learning tasks, and was not due to the research demand. Students practiced their English speaking ability and pronunciation by completing seven tasks with different speaking contents. The goal for the students was to stay on the tasks and as they improved their process of speaking practices using the app.

In the first week, the participants were asked to download and install the apps depending on which mobile device they used. Both apps generate immediate results based on the students' spoken words that allow them to compare the output to the text that they just spoke. In previous years, other STR apps such as Google Docs voice typing and Evernote were used with students and found to be inadequate or difficult to use. Thus, the need to find more appropriate apps resulting in using Speechy and SpeechNotes. The majority of students found that these apps were easy to use and were able to assist them to improve on the accuracy of their oral reading. Reflections from students indicated that they are more user friendly and overall better apps to use.

In class, the oral English training tasks lasted for 50 min each time. At the beginning of each task, students were instructed to read the full text of a passage from a TOEFL exam preparation book by Edmunds et al., (2009) in detail, and find and encircle unfamiliar words. Then, the instructor guided the students to read and understand the passage. Students were reminded to pay attention to the word and sentence pronunciation and intonation. Next, students were given 20 min to practice on their own before reading aloud the passage for 2 min on the app. Students may use the STR app to perform oral reading twice. They were then asked to choose, take a screenshot, and upload the results with fewer erroneously transcribed words on the TronClass, the learning management system provided by the university (Fig. 1). Students also have to take a photo and upload the passage script after the erroneously transcribed words were identified (Fig. 2). The procedure of each STR app task is as follows:

Fig. 1 Screenshot of STR app result

> **Note-3**
>
> Embroidery is today's topic what is it embroidery is needlework in which designs and features are created by stitching strands of some material to a layer of another material malt embroidery you say strength of friends or what embroidery could use different material why are or later stripped embroidery are you sure he s straight woven fabric such as spincast or comments a five can be in Boyertown too many materials things later felt and plastic embroidery has traditionally been used red clothing and household furnishings such as tablecloths towels and bedding
>
> But you can embroider anything as long as it is made from an evenly woman Fabric and can be held firmly in the hand spatial embroidery hoop embroidery has also been used as an art fall decoration fall decoration the most famous imported artwork is called the Bayeux Tapestry b a r i e u x it was made in England in the air 1077 live camera Straight A's like of huge van cat 870 M long and 0.5 meters wide it has 50 Isaias last tell the story of the events leading to the Battle of Hastings in 1066

Traditional EFL courses are usually structured with specific tasks, using specific materials and to be completed at the teacher's pace. By using the apps, students are allowed to complete tasks that they find more interesting and in their own time as well as receive immediate feedback. Furthermore, students can find passages relating to topics are more authentic and useful while preparing for the standardized English tests. By freeing the students from the teacher's constraints, it allows the students to become less anxious, more self-directed and purse topics that are of more interest to them.

Embroidery is today's topic. What is it? Embroidery is needlework in which designs and pictures are created by stitching strands of some material onto a layer of another material. Most embroidery uses strands of thread or wool, but embroidery could use different material, such as wire or leather strands. Embroidery designs are usually stitched into a woven fabric, such as blankets or carpets, but designs can be embroidered onto many materials, including leather, felt, and plastic. Embroidery has traditionally been used to decorate clothing and household furnishings such as tablecloths, towels and bedding.
But you can embroider anything as long as it is made from an evenly woven fabric and can be held firmly in the hand or in a special embroidery hoop. Embroidery has also been used as an art form and for decoration. One of the most famous embroidered artworks is called the Bayeux Tapestry - B-A-Y-E-U-X. It was made in England in the year 1077. This tapestry is like a huge blanket. It's 70 meters long and 0.5 meters wide. It has 58 scenes that tell the story of the events leading to the Battle of Hastings in 1066.

Fig. 2 Sample passage script of STR app task

3.3 Instrument and data analysis

In order to measure the effectiveness of repetitive training on spoken English and pronunciation through STR app tasks, several data collection methods were used which include pre- and posttests, STR app tasks results, field-observation notes, student reflective journal, and end of class survey. First of all, the matched tests from the speaking section of the mock up TOEFL iBT test were adopted from the Barron's TOEFL preparation book by Sharpe (2006). In the first week, the pre-test was administered using one of the mock up tests. From the second week to the end of semester, the course was taught accordingly. Then in the final week, a matched mock up test from the same book was used to obtain the post-test results. Both tests consisted of six speaking tasks and the test time is approximately 20 min. The test results were evaluated by two trained research colleagues and analyzed using paired sample t-test. The inter-rater reliability between two raters based on Gwet (2014) revealed that the study's intra-class correlation coefficient (ICC) were 0.937 with 95% CI [0.862, 0.971] on the speaking section of TOEFL iBT tests for the pre-tests, and 0.842 with 95% CI [0.653, 0.928] for the post-test. These indicated that both pre- and posttests had excellent reliability.

To calculate the effectiveness of this study's STR app tasks, the total number of erroneously transcribed words in the task results from each student was compared to determine if the number of errors made by each student decreased over time. The accuracy of STR app task results among students were calculated to examine whether it increased after they completed the tasks. Further, the number of erroneously transcribed words in the assigned passage was counted; thus, it serves as numerators of ratios whose denominators are the total numbers in assigned passages. For example, a student might produce 200 words in the assigned passage, 50 of which are incorrect; therefore, the student's accuracy is 75%. The accuracy rate is calculated using Eq. 1.

$$\frac{total\ number\ of\ words - -total\ number\ of\ erroneously\ transcribed\ words}{total\ number\ of\ words} \times 100 = accuracy\ rate$$

(Eq. 1)

Field-observation notes were analyzed using quantitative content analysis. The end of class survey includes items that consider the effectiveness and the perceptions of students on the STR app tasks for EFL speaking training. It contained 15 closed-ended items with 4 open-ended questions to help understand students' perspectives of the STR app tasks. The quantitative data obtained from the survey were analyzed with descriptive statistics using SPSS. The qualitative data obtained from the survey and student reflective journal were analyzed using conceptually clustered matrices. Finally, triangulation was used to ensure the validity and reliability of the measurement.

4 Results

4.1 Pre- and post-tests

As mentioned above, the pre- and post-test were mock up TOFEL tests, the results shown are paired sample t-tests which were performed to determine whether there is a difference between the students' speaking ability before and after implementing STR app tasks. In total, there are six speaking tasks with scores ranging from 0 to 30 (Good: 26–30, Fair: 18–25, Limited: 10–17, Weak: 0–9). In terms of students' speaking ability, the results demonstrated that students have fair speaking ability. As illustrated in Table 1, the difference between the scaled speaking scores of the pre- and post-test was significant ($t = -4.058$, $p > 0.05$), indicating that students' speaking ability improved after the implementation of STR app tasks. It was also found that all students' posttest scores ($M = 20.07$, $SD = 4.150$) were significantly higher than their pretest scores ($M = 18.33$, $SD = 5.650$). This means that compared to the conventional EFL speaking training course, the use of STR mobile learning app task was more beneficial in increasing students' accuracy rate of English words in terms of speaking. As there are no previous studies using pre- and post-data that has measured the improvement of speaking has been published, this become the impetus for completing this study.

4.2 Practice results of oral english training

Each student was required to complete the 7 STR app tasks in class and submit the results to TronClass for over a semester. The accuracy rates of English words in the context were calculated to analyze students' speaking practice results as detailed in Table 2. The average accuracy rates of tasks increased toward the end of the study period (Fig. 3). In addition, t-test results revealed that there were statistically signifi-

Table 1 Results of paired sample t-tests of the pre- and posttest		N	Mean	SD	*t*	*Sig.*
	Pretest	27	18.33	5.650	-4.058	0.000*
	Posttest	27	20.07	4.150		
	*P < 0.05					

cant differences in the accuracy rate of English words across the 7 tasks. These results indicated that the decrease in the students' erroneously transcribed words demonstrated an increase in their accuracy rates across the 7 STR app tasks.

4.3 Classroom observations

During the study period, this study conducted 7 classroom observations while the STR app tasks in class were being held. In the first few weeks, students generally had a positive attitude toward the tasks. It was observed that they were eager to find a comfortable spot in the classroom to start their own practices and get ready for the formal recordings. When the results were not good enough, they kept trying and repeating the words until they get it right. Some students would come to ask for the instructor's help during the self-practice period, while others talk to their classmates to verify the correctness of their pronunciation. Even after the tasks, some of them would still continue talking about the tasks. One thing that bothered students with iOS phones was the limitation of the Speech app; it stops after 60 s. Thus, the researcher extended the total recoding time to 120 s to allow the students to start a new recording for the rest of the speaking context. Occasionally, students mentioned the app missed few of the words they spoke. Toward the end of the study period, few students complained that the app failed to show the words promptly or correctly and sometimes missed many words. Since this only happened to a few students, the

Table 2 Statistical analysis of the accuracy rate of students' speaking practice results

	N	M	SD	t	df	Sig.
Task 1	24	31.75	37.076	9.018	23	0.000*
Task 2	23	48.00	23.004	10.841	22	0.000*
Task 3	23	54.26	32.927	6.808	22	0.000*
Task 4	23	61.26	17.158	10.828	22	0.000*
Task 5	25	66.52	12.474	13.420	24	0.000*
Task 6	24	60.71	20.224	9.518	23	0.000*
Task 7	22	67.55	16.203	9.395	21	0.000*

*P<0.05

Fig. 3 The accuracy rate across STR mobile learning app tasks
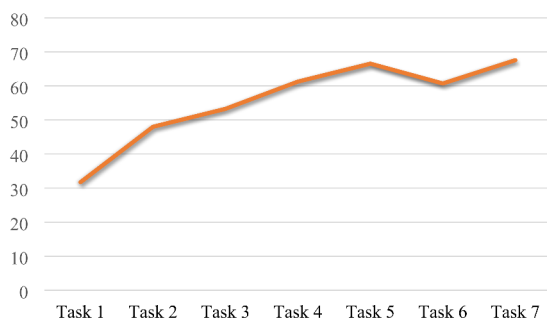
malfunction might have been caused by the cell phone unit or the Internet connection, and not the app itself.

4.4 Students' reflective journal

To encourage the students to elaborate their views on the STR app tasks, students were asked to post their reflective comments on the Facebook class page after each task. This assignment was set up to better understand not only the effect of the tasks, but also the students' experiences with the STR app. After data collection, conceptually clustered matrices developed by Miles & Huberman (1994) were used to reduce and present data, and draw and verify the conclusion. This allowed the related reflections to be put together so that a list of various Facebook postings made by the students can be sorted and coded. By analyzing student reflective journals, their insights about the app were identified and were summarized to evaluate the effect of the STR app tasks. The opinions were arranged from most to least frequently mentioned and coded in three categories: (1) learning achievement, (2) perception, and (3) engagement.

As shown in Table 3, the most frequently mentioned opinion belonged in the perception category, which revealed that students believed that the tasks were an

Table 3 Students' insights on the STR app tasks

Students' reflective journals	Code
STR app task is an interesting way to practice speaking in English.	Perception
I feel my speaking ability improved over time.	Learning achievement
The task is kind of difficult, but it helped improve my speaking ability.	Perception
My English speaking ability has become more fluent.	Learning achievement
The STR app task is motivating.	Perception
The tasks made me realize that I am not good enough in speaking English, especially with word pronunciation.	Perception
I am aware of my weaknesses in English speaking and the need to practice more.	Learning achievement
It was more engaging compared to the conventional teaching methods.	Engagement
The context of the tasks widened my vocabulary.	Learning achievement
The tasks helped me learn to concentrate better on the English oral practices.	Engagement
I feel more confident in speaking English now.	Perception
The STR app tasks provided a really meaningful way to practice oral English.	Perception
I am excited about the tasks and I am always looking forward to the next one.	Perception
It is frustrating that the app missed some words of the words that I spoke.	Perception

interesting way of practicing their English speaking. For the learning achievement category, most students mentioned that their speaking ability improved and became more fluent because of the tasks. The tasks helped them identify their weaknesses which made them realized consistent practice is needed to improve their speaking ability. Furthermore, students expressed that they were more engaged in class when the STR app tasks were employed than when the teaching methods in traditional EFL classes were implemented. The tasks also made them learn to focus their attention on the speaking context which helped them learn better. In terms of negative viewpoints, one issue was occasionally mentioned: ease of use. Students stated that "It was sort of annoying when the app did not function properly and the words I speak did not show up on the screen."

4.5 End of class survey

In the final class, the students were asked to answer a survey questionnaire consisting of 15 close ended questions using a five-point Likert scale (1 for strongly disagree and 5 for strongly agree). The Cronbach's alpha score was 0.939 which demonstrates excellent reliability. As in most cases, a positive response is interpreted when the mean score of the item is 3.40 or higher. As shown in Table 4, the results of the survey indicated that students generally have positive views on the STR app tasks with an overall mean score of 4.24 (SD = 0.57624). The students' positive attitudes were evident within the learning achievement (M = 4.29), perception (M = 4.25) and engagement (M = 4.15) domains. Among the students, 83.7% of them had highly positive (45.9%) or positive (37.8%) scores on the learning achievement category of STR mobile learning app tasks. These students reported that their English pronunciation, accent, and intonation improved significantly and that they are now aware of how to effectively improve their speaking ability. Within the perception category, most of the students (44.2%) agreed and strongly agreed (42.1%) that STR app tasks were very easy to operate and were highly motivating, interesting, and an effective method to practice speaking in English. The tasks provided meaningful learning experience and were better than the traditional classroom teaching methods. Further, most mentioned that they will recommend the app to other learners. For the engagement category, the majority of the respondents conveyed positive (43.3%) and highly positive (38.2%) engagements in terms of concerting on the STR app tasks, learning through it, and continuing to use it to improve their speaking ability in the near future.

Table 4 Distribution of mean scores on the attitude toward SRT app tasks

Constructs	Percentage					M	SD
	SD	D	N	A	SA		
Learning Achievement	0.0	0.7	15.6	37.8	45.9	4.29	0.58791
Perception	0.5	2.1	11.1	44.2	42.1	4.25	0.55726
Engagement	0.0	4.9	13.6	43.3	38.2	4.15	0.71213
Overall attitude	0.2	2.6	13.4	41.7	42.1	4.24	0.57624

Note. (1) SD, strongly disagree; (2) D, disagree; (3) N, neutral; (4) A, agree; (5) SA, strongly agree

The students were also asked to answer the open-ended questions which provided a deeper understanding of the results generated from the quantitative data analysis. In order to triangulate, data obtained from the open-ended questions were analyzed

Table 5 Summary of responses on open-ended question with constructs and memos

Questions/ Students' Responses	Constructs	Memo
Q To what extend does the STR app tasks helped you in speaking in English?	Learning achievement	Since speaking in English in the STR app tasks is unavoidable every week, students eventually learned to self-correct their mistakes and improve their oral skills.
SR The app will tell me if I pronounced the texts correctly. My accuracy rates greatly improved. The tasks really helped me fine-tune my pronunciation. (Number of students who made similar comments: 10)		
I can speak more coherently and with a logical flow now. (Number of students who made similar comments: 9)		
I am more confident in speaking in English. (Number of students who made similar comments: 5)		
Q What were the difficulties or challenges you encountered in the STR app tasks?	Perception	Students were notified that the STR app is not 100% accurate in converting the speech to text during the class orientation. With only small imprecisions, it is still a great tool for practicing speaking in English.
SR • In order for the app to convert the spoken word into text correctly, I need to have precise pronunciation. (Number of students who made similar comments: 11) • It was frustrating when the app cannot recognize my voice, even though I am pretty sure that my pronunciations were correct. (Number of students who made similar comments: 6)		
Q What is your perception of these tasks?	Perception	Even though the reading texts in the tasks were difficult for them, as mentioned by two students, most still tried their best to complete the tasks and believed that the tasks were very useful for practicing speaking in English.
SR Although I need to put in a great deal of effort, I believe it was absolutely worth trying. (Number of students who made similar comments: 13) The STR app tasks pushed me to work toward a better result. (Number of students who made similar comments: 9) I am more aware of what I need to do to improve my English speaking ability (Number of students who made similar comments: 6)		
Q Is there anything else you would like to say about this course?	Engagement	Two students further expressed that they love to come to this class particularly because the tasks make the learning very effective and not boring. Even though obtaining low scores in the STR app tasks can be frustrating, learning one's weakness in speaking can result in considerable improvements.
SR It is a more meaningful way to learn which makes me want to attend the class. (Number of students who made similar comments: 7) I will continue using this app to practice my English speaking ability. (Number of students who made similar comments: 5) I will recommend it to others who are interested in learning English. (Number of students who made similar comments: 4)		

Note: Q=question; SR=sample responses from students

using similar methods as those used for the students' reflective posts. Matrices were also created to facilitate the coding and categorization process during data analysis. Three constructs were then adopted which are summarized in Table 5.

4.6 Results of triangulation

To add depth to the data and enhance the validity and reliability of the results, classroom observations, student reflective journal, and end of class survey about the implementation of STR app tasks were triangulated as shown in Table 6.

Table 6 Comparison of findings from a variety of data types regarding the use of STR app tasks

The effects of STR app tasks	Sample evidences		
	Classroom observations	Student reflective journal	End of class survey
Improved English speaking ability	Most students improved their English speaking ability judging by the accuracy rate of their English words.	My English speaking ability is getting better and I am becoming more fluent.	My English pronunciation have improved greatly by practicing through the STR app.
Enjoyed the learning process	Students would practice as many times as they can to complete the tasks.	It is an interesting and motivating way to practice speaking in English. I learned to concentrate better during practices.	It was a pleasant and meaningful experience to learn English through the tasks.
Felt confident in speaking	Students were becoming increasingly confident toward the end of the semester.	I feel that I am much more confident in speaking in English.	I am not that afraid of speaking in English any more.
Motivated to come to class	The attendance rate was high. Students hardly missed any class.	I love attending this class more than other conventional ones. The tasks were challenging, but effective.	This is the type of class that I feel I can be fully involved in the learning process.
Increased learner engagement	During the tasks, all students were actively participating and some of them showed eagerness to use the app.	I find myself more engaged in learning than when conventional teaching methods were used.	I feel that I can concentrate better on learning when working on the tasks.

5 Discussion

In this study, both quantitative and qualitative methods were used to gather data.

The pre- and post-test results indicated that the difference in students' speaking ability before and after implementing the STR app tasks were statistically significant. The speaking practice results confirmed that the students' average accuracy rates were significantly higher toward the end of the semester. The classroom observation results also supported these results. Based on the classroom observations results, students generally perceived the STR app as an effective tool for their oral English training. During the class sessions, all students remained focused on the tasks and tried to practice as many times as possible before the recordings. The analysis of students' reflective journals revealed that the STR app tasks did not only motivate them to do well in their oral speaking training, but also helped them improve their English speaking ability, particularly their coherence, pronunciation, and fluency.

In addition, the results of the end of class survey confirmed the effectiveness of the STR app tasks. Most students (93%) claimed that that their English pronunciation, accent, and intonation greatly improved. As local language influences the way students learn to speak English, especially in pronunciation and intonation (Krishna & Krishnan, 2014), but this study's participants stated that using STR apps in the classrooms has assisted them to improve their pronunciation, intonation and accent to be more like native English speakers. Therefore, it is evident that using STR apps are beneficial as one of the teaching tools in an ESL classroom. Moreover, several students mentioned that with the STR app tasks, students became aware of how to effectively improve their English speaking ability. The students also claimed that the app is very easy to use and the tasks are interesting. It made them feel more confident and motivated to attend their English classes. Furthermore, most students mentioned that the tasks enabled them to realize that learning is meaningful and that the app helped them become more engaged in learning in class. They preferred to use the STR app tasks to enhance their oral speaking ability over conventional methods.

One difficulty that bothered a small number of students in the learning process was the inaccuracy of the STR app results. Few students mentioned that sometimes the speech recognition did not function correctly. It failed to display the words promptly on the screen, suggested wrong words, or missed some spoken words. Although the STR technology is very useful, it is not yet perfect. This is why the present research explained to the students during their class orientation that they may experience some glitches and problems in the app. During the orientation, it was mentioned that the STR app might not be able to distinguish homonyms, slang, technical words, and acronyms. In addition, background noise interference and bad internet connection might also cause problems. However, the students were told not to worry about getting all the words correctly, and that an accuracy rate of 80% is already satisfactory.

6 Conclusions

In the present study, the STR app tasks were implemented to an EFL oral speaking training course in a university to explore their effects on the students' speaking profi-

ciency. There were two main findings obtained through this study: First, the statistical results indicated that the STR app helped increase students' accuracy rate of English spoken words and their TOEFL speaking scores, which is an important indicator of students' improvement of speaking achievement. Studies have also noted that STR technology has reached a satisfactory level of accuracy (Carrier, 2017; Huang et al., 2016) and it has been widely recognized as a crucial factor that contributes to the success of language learning. The students seemed to have totally accepted the idea of using STR app into the EFL oral speaking training course and were able to immerse themselves in the tasks. All students believed that the STR app is an effective educational tool that can bring significant improvements in their EFL oral training. Second, the results of the qualitative analyses indicated that the students had a strong positive attitude toward the use of the STR app in their EFL oral speaking training. In their reflective journals, the students mentioned that the STR app tasks were useful in improving their English speaking ability. The tasks did not only widen the students' vocabulary, but also made them speak in English more fluently. These results corresponded with Ahn & Lee's study in 2016 which concluded that students gave positive comments that speech recognition technology is useful and believed that using it was beneficial to improving their EFL speaking.

This study has several contributions to the knowledge gap and literature in the field of STR technology and MALL for academic learning. It presented several empirical evidences that validated the effectiveness of the STR app tasks on EFL students' oral speaking performance in class and speaking ability in general. This can be seen on the improvement of students' weekly average accuracy rates and their pre- and post-tests results. Further, students' positive perceptions toward the benefits of the STR app tasks on EFL oral speaking training were demonstrated. Particularly on the improvement of their English pronunciation, accent and intonation that assisted them to become a more confident and fluent speaker. Moreover, it confirmed that students' engagement in learning during their class significantly improved with the implementation of the STR app tasks. The results of classroom observations, reflective journals, and end of class survey all showed that students enjoyed the learning process and were moted to come to class when the STR app tasks were implemented. The students felt that they can concentrate better, engage more and participate actively. Lastly, implications and recommendations with respect to further practices in the field were made for EFL teachers and students in Taiwan as well as in the large global community.

Despite trying its best to provide an objective and comprehensive study, this research still has limitations. The sample size in this study was relatively small; future research should recruit a larger number of students to improve the generalizability of the results of the STR app implementation. Also, this study implemented the STR app for one semester only. It is predicted that a year-long study might generate different results. Future research may consider extending this research topic in several ways. First, the effectiveness of the STR app tasks on EFL oral speaking training can be investigated by qualitative methods such as a case study, to gather more in-depth insights of the students' learning process. Second, students' background characteristics such as learning style, English proficiency, and gender were not explored in this study; thus, future research may focus on these factors because they could affect

the effectiveness of the STR app tasks. As noted by Lai & Zheng (2018), learners' characteristics are primary considerations when implementing mobile technology for language learning. Third, the implementation of STR technology to facilitate EFL learning can be extended by studying its impact on teachers. For instance, future studies may investigate how teachers could apply the STR technology for a variety of class activities or tasks to assist in teaching and testing their students' English speaking ability.

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. https://doi.org/10.1111/bjet.12354

Bian, F. (2013). The influence of Chinese stress on English pronunciation teaching and learning. *English Language Teaching*, 6(11), 199–211

Carrier, M. (2017). Automated speech recognition in language learning: Potential models, benefits and impact. *Training Language and Culture*, 1, 46–61. https://doi.org/10.29366/2017tlc.1.1.3

Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2), 624–645. https://doi.org/10.1016/j.compedu.2007.06.011

Chen, C. M., Liu, H., & Huang, H. B. (2019). Effects of a mobile game-based English vocabulary learning app on learners' perceptions and learning performance: A case study of Taiwanese EFL learners. *ReCALL*, 31(2), 170–188. https://doi.org/10.1017/S0958344018000228

Chen, H. H. J. (2011). Developing and evaluating an oral skills training website supported by automatic speech recognition technology. *ReCALL, 23*(1), 59–78. https://doi.org/10.1017/S0958344010000285Chen, T. H., & Lin, C. C. (2018). Enhancing L2 English Learning through Mobile-Assisted TBLT: EFL Learners' Perspectives. *Journal of Asia TEFL, 15*(2), 453–461

Chen, Y. L., & Hsu, C. C. (2020). Self-regulated mobile game-based English learning in a virtual reality environment. *Computers & Education*, 154, 103910. https://doi.org/10.1016/j.compedu.2020.103910

Coniam, D. (1998). The use of speech recognition software as an English language oral assessment instrument: an exploratory study. *CALICO Journal*, 15(4), 7–23. https://doi.org/10.1558/cj.v15i4.7-23

Cox, T. L., & Davies, R. S. (2012). Using automatic speech recognition technology with elicited oral response testing. *CALICO Journal*, 29(4), 601–618. https://doi.org/10.1558/cj.29.4.601-618

Edmunds, P., McKinnon, N., & Zeter, J. (2009). *Developing skills for the TOEFL iBT* (2,2nd, ed.). Compass Publishing

Elliott, J. (1991). *Action research for educational change*. Philadelphia, PA: Open University Press

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning and Technology*, 15(2), 2–11

Gromik, N. A. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, 58(1), 223–230. https://doi.org/10.1016/j.compedu.2011.06.013

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The Definitive guide to measuring the extent of agreement among raters* (4nd ed.). Gaithersburg, USA: Advanced Analytics, LLC

Hansen, J. H. L., & Bořil, H. (2018). On the issues of intra-speaker variability and realism in speech, speaker, and language recognition tasks. *Speech Communication*, 101, 94–108. https://doi.org/10.1016/j.specom.2018.05.004

Heil, C. R., Wu, J. S., Lee, J. J., & Schmidt, T. (2016). A review of mobile language learning applications: Trends, challenges, and opportunities. *The EuroCALL Review*, 24(2), 32–50. https://doi.org/10.4995/eurocall.2016.6402

Hsu, C. K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL: the Journal of EUROCALL*, 27(1), 84–103. https://doi.org/10.1017/S0958344014000214

Hsu, C. K., Hwang, G. J., & Chang, C. K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, 63, 327–336. https://doi.org/10.1016/j.compedu.2012.12.004

Hsu, L. (2013). English as a foreign language learners' perception of mobile assisted language learning: a cross-national study. *Computer Assisted Language Learning*, 26(3), 197–213. https://doi.org/10.1080/09588221.2011.649485

Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, 58(1), 273–282. https://doi.org/10.1016/j.compedu.2011.08.008

Huang, Y. M., Shadiev, R., & Hwang, W. Y. (2016). Investigating the effectiveness of speech-to-text recognition applications on learning performance and cognitive load. *Computers & Education*, 101, 15–28. https://doi.org/10.1016/j.compedu.2016.05.011

Jamaldeen, F., Hewagamage, P., & Ekanayaka, Y. (2018). Design guidelines for creating mobile language learning applications. *International Journal of Interactive Mobile Technologies (iJIM)*, 12, 52. https://doi.org/10.3991/ijim.v12i3.8153

Jia, J., Chen, Y., Ding, Z., & Ruan, M. (2012). Effects of a vocabulary acquisition and assessment system on students' performance in a blended learning class for English subject. *Computers & Education*, 58(1), 63–76. https://doi.org/10.1016/j.compedu.2011.08.002

Krishna, G. R., & Krishnan, R. (2014). Influence of mother rongue on English accent. *Proceedings of the 11th International conference on natural language processing*, India, 63–67

Kukulska-Hulme, A. (2012). How should the higher education workforce adapt to advancements in technology for teaching and learning? *The Internet and Higher Education*, 15(4), 247–254. https://doi.org/10.1016/j.iheduc.2011.12.002

Kung, F. W. (2017). Teaching and learning English as a foreign language in Taiwan: A socio-cultural analysis. *TESL-EJ*, 21(2), 1–15

Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, 30(3), 299–318. https://doi.org/10.1017/S0958344017000258

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). London, UK: SAGE Publications

Niño, A. (2015). Language learners perceptions and experiences on the use of mobile applications for independent language learning in higher education. *The IAFOR Journal of Education*, 3(SE), 73–84

Pal, S. (2013). Mother tongue influence on spoken English. *Conference proceedings ICT for language learning*, Italy.454–458

Persson, V., & Nouri, J. (2018). A systematic review of second language learning with mobile technologies. *iJET*, 13(2), 188–210

Rosell-Aguilar, F. (2017). State of the app: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal*, 34(2), 243–258. https://doi.org/10.1558/cj.27623

Shadiev, R., Hwang, W. Y., Chen, N. S., & Huang, Y. M. (2014). Review of speech-to-text recognition technology for enhancing learning. *Educational Technology & Society*, 17(4), 65–84

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. https://doi.org/10.1016/j.compedu.2016.03.013

Shadiev, R., Huang, Y. M., & Hwang, J. P. (2017a). Investigating the effectiveness of speech-to-text recognition applications on learning performance, attention, and meditation. 65(5),1239–1261. https://doi.org/10.1007/s11423-017-9516-3

Shadiev, R., Wu, T., & Huang, Y. M. (2017b). Enhancing learning performance, attention, and meditation using a speech-to-text recognition application: evidence from multiple data sources. Interactive Learning Environments, 25, 249–261

Sharpe, P. J. (2006). *Barron's TOEFL iBT: Internet-based test* (12th ed.). Hauppauge, NY: Barron's

Strik, H., Cornillie, F., Colpaert, J., Doremalen, J., & Cucchiarini, C. (2009). Developing a CALL system for practicing oral proficiency: How to design for speech technology, pedagogy and learners. *Speech Communication*. Retrieved from http://hstrik.ruhosting.nl/wordpress/wp-content/uploads/2013/03/a153-DISCO_Design-SLaTE09.pdf

Wankel, L. A., & Blessinger, P. (2013). New pathways in higher education: An introduction to using mobile technologies. In A. W. Laura, & B. Patrick (Eds.), *Increasing Student engagement and retention using mobile applications: Smartphones, skype and texting technologies* (6 vol., pp. 3–17). Emerald Group Publishing Limited. Part D

Ya, F. Y., Ching, J. C., & Chih, K. C. (2013). Mobile assisted language learning: Overview of literature from 2008 to 2012. In *The 21st International Conference on Computers in Education* (ICCE 2013), Denpasar Bali, Indonesia

Zhang, Y., & Liu, L. (2018). Using Computer Speech Recognition Technology to Evaluate Spoken English. *Educational Sciences: Theory & Practice*, 18(5), 1341–1350. https://doi.org/10.12738/estp.2018.5.033

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Kate Tzu-Ching Chen Kate Tzu-Ching Chen has an Ed.D. in Curriculum and Instruction from the University of South Dakota, USA. She is currently engaging in researches on educational technology. She presently resides in Taiwan and teaches as the Chairperson and an Associate Professor at Chaoyang University of Technology, the Department of Applied English.