Exploring the Effects of Using Automatic Speech Recognition on EFL University Students with High Speaking Anxiety

Wenyi Li[1,*], Maslawati Mohamad[1], and Huay Woon You[2]

[1]Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[2]Pusat PERMATA@Pintar Negara, Universiti Kebangsaan Malaysia, Selangor, Malaysia
Email: p121641@siswa.ukm.edu.my (W.Y.L.); maslawati@ukm.edu.my (M.M.); hwyou@ukm.edu.my (H.W.Y.)
Manuscript received October 15, 2024; revised November 1, 2024; accepted November 20, 2024; published January 20, 2025

***Abstract*—This study explores the effects of Automatic Speech Recognition (ASR) technology on English as a Foreign Language (EFL) students with high speaking anxiety. Conducted among 98 first-year university students in China, the research employed a mixed-methods approach, integrating quantitative data from a speaking anxiety questionnaire with qualitative insights from students' reflective journals. Over 14 weeks, participants engaged in ASR-based speaking tasks, receiving real-time feedback to improve pronunciation and speaking competence. Findings indicate that ASR significantly reduced speaking anxiety, particularly in unprepared speaking tasks, and increased learners' confidence in speaking English. However, peer-related anxieties persisted, suggesting ASR's limitations in alleviating social anxieties in group contexts. The study concludes that while ASR is an effective tool for individual language practice, it should be supplemented with peer interaction strategies to address speaking anxiety more comprehensively.**

***Keywords*—automatic speech recognition, English as a Foreign Language (EFL) learners, speaking anxiety, mixed-methods, language learning**

I. Introduction

In the modern era, rapid advancements in science and technology have revolutionized various industries, and education is no exception. Information technology has become deeply embedded in language learning, transforming traditional methods and offering new opportunities for enhancing learners' skills [1]. One such innovation is Automatic Speech Recognition (ASR), which decodes and transcribes spoken language through pattern analysis, modeling, and algorithms [2]. With its ability to provide immediate and personalized feedback on pronunciation, ASR has gained significant popularity in English as a Foreign Language (EFL) classrooms [3]. This technology enables learners to practice their speaking skills anytime and anywhere, breaking traditional limitations of time and space in language education [4].

Given the flexibility and accessibility of ASR, it has become an indispensable tool in supporting speaking competence for EFL learners. Research has shown that ASR simplifies language learning by offering instant feedback, which helps learners improve their pronunciation and fluency [5]. Popular applications such as Google Web Speech, iFlytek Voice Input, and Speechnotes are increasingly used to enhance speaking competence [6]. The key strength of ASR lies in its ability to provide a private, low-pressure environment for language practice. It allows learners to practice language skills independently outside formal settings [7], enabling them to improve their speaking abilities without the immediate pressure of peer or teacher evaluation.

Another significant benefit of ASR is that it addresses the challenge of limited feedback in traditional classroom settings. Effective language learning requires constant feedback, but learners often struggle to monitor their speech because of interference from their native language [8]. Since it is impractical for teachers to provide immediate, personalized feedback to every student, ASR fills this crucial gap by offering instant corrections and suggestions, enabling learners to practice more efficiently and independently [9].

A persistent challenge that EFL learners face when practicing speaking is anxiety. Speaking anxiety can be described as tension, nervousness, or worry, especially when speaking a foreign language [10]. In the context of learning English, speaking anxiety manifests as the fear or discomfort that arises when learners are required to speak or participate in conversations [11]. The anxiety tends to be more pronounced when learners must perform in front of others, such as during class activities or unprepared presentations [12].

In recent years, an increasing number of longitudinal studies have examined speaking anxiety, often using well-established scales to measure anxiety levels among learners [13]. While these standardized scales have proven reliable, there are concerns about the transient nature of the data they produce, as speaking anxiety can fluctuate based on various factors [14]. Given language learning anxiety's complex and dynamic nature, relying on quantitative methods may limit understanding of this phenomenon. Zhao's comparative study of 198 research papers highlights the prevalence of quantitative methods in China [15], while advocating advocates for more qualitative approaches to complement existing research.

Furthermore, an analysis of journals from the China National Knowledge Infrastructure (CNKI) between 2008 and 2024 reveals a significant gap: out of 48 studies on English-speaking anxiety among Chinese university students, only three focused on English majors [16]. This gap is particularly significant given that many English majors are likely to pursue careers in English education and translation, fields critical to China's future development in global communication. Therefore, understanding and addressing the speaking anxiety of English majors is essential to help them succeed in these careers and contribute more effectively to English education in China.

Given these identified gaps in current research and the potential of ASR technology, this study seeks to address the following specific research questions:

1) To what extent does ASR technology reduce speaking anxiety among Chinese English majors?

2) How do students perceive and experience ASR as a tool for reducing speaking anxiety?

Based on the first research question, the following hypothesis is proposed:

Hypothesis: ASR technology significantly reduces the overall speaking anxiety of Chinese English majors.

This research adopted a mixed-methods approach, combining quantitative and qualitative analyses to comprehensively evaluate the effects of ASR technology on speaking anxiety. The findings indicate that ASR positively reduces speaking anxiety. Students also expressed favorable attitudes toward ASR-based learning, suggesting that this technology holds great potential as a tool for language learning. As language educators continue to explore new technologies, it is crucial to consider learners' unique preferences and needs to integrate ASR into pedagogical practices effectively.

II. Literature Review

A. Automatic Speech Recognition (ASR) in the EFL Context

Automatic Speech Recognition (ASR) technology enables machines to decode and transcribe spoken language independently. It has become integral to modern communication tools, such as Siri, and various speech dictation applications [17]. As a specialized application of artificial intelligence in natural language processing, ASR is increasingly used in communication, particularly in the EFL context [18]. Its efficacy in enhancing language learning experiences has been well-documented, where it supports pronunciation practice and speaking assessments [19]. Unlike traditional classroom methodologies, often limited by time and resources, ASR offers unique advantages by providing extensive practice opportunities and consistent, unbiased feedback, which are challenging to achieve in conventional settings [20].

Extensive empirical research supports the application of ASR in EFL education, particularly in pronunciation practice and speaking assessment [9, 19, 21]. For pronunciation practice, ASR tools like ELSA Speak provide immediate feedback on pronunciation accuracy, helping learners monitor their progress and refine their speaking skills [22]. Studies show that such applications can effectively enhance speaking competence [23]. For speaking assessment, ASR provides a more objective and standardized alternative to traditional human scoring. For example, the TOEFL IBT Test uses ASR to evaluate spoken responses based on pronunciation, fluency, and intonation [24]. This automated approach promotes fairness and reduces biases, enhancing language evaluation's reliability.

Nevertheless, many ASR programs require learners to practice predetermined words and phrases, limiting opportunities for communicative practice and hindering learner autonomy [25]. Moreover, prior research has predominantly focused on ASR-based computer-assisted pronunciation training, with comparatively less attention given to mobile dictation ASR in EFL instruction [5]. Given that dictation ASR applications are designed for native speakers and accommodate natural language use, they hold the potential for advanced speaking practice, enabling EFL learners to self-monitor their spoken output [21]. Therefore, this study aims to explore the influence of ASR on English-speaking anxiety by using dictation application software to assist in English-speaking learning.

B. Speaking Anxiety in the EFL Context

Anxiety is a distressing emotional state that stems from physiological arousal. It often involves tension, apprehension, nervousness, and worry triggered by the autonomic nervous system activation [10, 26]. In foreign language learning, anxiety mainly appears when learners are required to speak [9]. Talking to a native speaker or a language teacher can lead to intense anxiety, reducing language learners' confidence to continue conversation [27]. Speaking tasks, such as live presentations or unprepared exercises, often heighten this anxiety, negatively impacting learners' speaking skills and self-confidence [12].

Recent research has focused on the impact of technology on speaking anxiety. Sho's research demonstrated that video conferencing platforms could reduce EFL learners' speaking anxiety by creating a less intimidating environment for language practice [28]. Similarly, Nguyen found that repeated speaking practice using mobile phone video recording tools lowered speaking anxiety among learners [29]. Bashori *et al.*'s study revealed that web-based language learning platforms, particularly those equipped with ASR technology, significantly reduced speaking anxiety by providing a low-pressure setting for practice [9]. Further supporting this trend, Rahman and Tomy demonstrated that intelligent personal assistants like Siri or Google Assistant significantly reduced EFL learners' speaking anxiety [30].

A careful review of existing studies on speaking anxiety shows that most employ mixed research methods. However, research on English-speaking anxiety in China predominantly relies on quantitative approaches, with limited integration of qualitative and quantitative analysis [15]. Additionally, according to CNKI, from 2008 to 2004, only 3 out of 48 studies on English anxiety among Chinese university students specifically focused on English majors [16]. This study seeks to address this gap by examining the speaking anxiety of English majors in Chinese universities, thereby enriching research in this field.

C. Identifying Research Gaps

Research on the application of ASR technology in mobile dictation software for EFL instruction still needs to be explored. Existing studies have primarily concentrated on ASR-based computer-assisted pronunciation training, often overlooking the potential of mobile dictation applications to enhance speaking practice through natural language use. These applications could offer EFL learners more significant opportunities for self-monitoring and engagement, promoting learner autonomy.

Furthermore, while a significant body of literature addresses speaking anxiety, much of it relies on quantitative methods, limiting understanding of its complexities, especially in the Chinese context. A review of recent studies indicates that only a tiny fraction specifically targets English majors despite their unique experiences and challenges with speaking anxiety. This study aims to address these gaps by investigating how mobile ASR-enabled dictation applications influence English-speaking anxiety among English majors in

Chinese universities, thereby contributing to a more comprehensive understanding of speaking anxiety in EFL settings.

III. Materials and Methods

This study employed a mixed-methods design, which integrates both quantitative and qualitative data to leverage the strengths of each approach. While quantitative data provide measurable and generalizable results, qualitative data offer deeper insights into participants' experiences and contexts [31]. By combining these methods, the study reinforced the reliability and validity of its findings through complementary data sources.

This study gathered quantitative data through a speaking anxiety questionnaire (SAQ), and qualitative data was obtained from learners' reflective journals. This combination allowed for a richer exploration of the effects of ASR technology on EFL learners' speaking anxiety. The quantitative data provided broader patterns and trends, while the qualitative data revealed more personal experiences and emotional responses. Consequently, these methods have facilitated a more nuanced and credible understanding of the phenomenon [32].

A. Participants

The present study was conducted in a public university in China. A total of 98 first-year English major undergraduates participated after providing consent. Each participant was drawn from three parallel EFL classes, all taught by the same instructor with 20 years of experience in the program. Each class had a weekly 80-minute face-to-face session on different weekdays for 14 weeks. During this time, participants used mobile ASR-enabled dictation applications to assist English-speaking learning. Table 1 below summarizes the demographic characteristics of the participants:

Table 1. Demographic characteristics of the participants

Demographic Variable	Category	Frequency	Percentage (%)
Gender	Female	77	78.6
	Male	21	21.4
Age (Years)	18	20	20.4
	19	38	38.8
	20	40	40.8
University Entrance Exam Score	Below 100	5	5.1
	100–120	20	20.4
	121–130	36	36.7
	131–140	25	25.5
	Above 140	12	12.2
Previous English Learning	6–8 years	19	19.4
	9–11 years	60	61.2
	12 years	19	19.4
Daily Time Using Mobile Devices to Learn English	Less than 30 minutes	20	20.4
	30–60 minutes	40	40.8
	61–90 minutes	25	25.5
	More than 90 minutes	13	13.3

Most participants were female (78.6%), with a smaller proportion of male participants (21.4%). In terms of age, the sample comprised students primarily aged 19 and 20, representing 38.8% and 40.8%. Regarding academic preparedness, the distribution of university entrance exam scores indicated that over half of the participants (74.4%) scored above 120, suggesting a competent level of English proficiency. Notably, most participants reported having 9-12 years of previous English learning experience (80.6%), which is expected to enhance their engagement with mobile ASR-enabled dictation applications. Data on daily usage of mobile devices for English learning indicated that most participants (79.6%) engaged with these resources for more than 30 minutes each day. This usage level suggests a strong possibility for integrating ASR applications into their language learning practices.

B. Data Collection Instruments

Data were collected through quantitative and qualitative tools, including a questionnaire and reflective journals. These instruments offered a comprehensive dataset that allowed for an in-depth analysis of ASR's effects on learners' speaking anxiety. A brief description of each instrument follows.

1) Questionnaire

The questionnaire (see Appendix Table A1) employed in this study was adapted from the Speaking anxiety questionnaire (SAQ) with a reliability score of 0.906 developed by Liu [33]. Liu's questionnaire is used as an instrument in this study because it provides a specific classification of items. Most previous studies classified the Foreign Language Classroom Anxiety Scale (FLCAS) items according to communication apprehension, fear of negative evaluation, and test anxiety or did not classify FLACS items [34]. Hovers, Horwitz acknowledges that FLACS, developed in 1986, does not provide a detailed classification [35]. She emphasized that the components of foreign language anxiety are not only communication apprehension, fear of negative evaluation, and test anxiety. Liu designed a questionnaire on anxiety in spoken English class by referring to the causes of speaking anxiety found in the existing literature [10, 36, 37] and selecting only the physical symptom items of anxiety, tension, and lack of confidence related to speaking anxiety in FLACS.

Table 2. The dimensions of the SAQ

Dimension	Total items (14)	Items No.	Source
Speaking English in class as self-confidence	4	1, 6,10, 11	Horwitz *et al.* (1986)
Making mistakes	2	2, 7	Horwitz *et al.* (1986)
Teacher input	2	3, 12	Horwitz *et al.* (1986)
Preparation	2	4, 14	Liu (2006)
Volunteering to speak	1	5	Dewaele & MacIntyre (2014)
Being singled out to speak English	1	8	Liu (2006)
Peer pressure	1	9	Dewaele & MacIntyre (2014)
Being laughed at	1	13	Dewaele & MacIntyre (2014)

Table 2 shows the dimensions of the SAQ. It has 14 items, covering such aspects related to speaking English in class as self-confidence (items1, 6, 10, 11), making mistakes (items 2, 7), teacher input (items 3, 12), preparation (items 4, 14), volunteering to speak (item 5), being singled out to speaking English (item 8), peer pressure (item 9), and be laughed at (item 13). SAQ is designed on a 5-point Likert-type ranging from "strongly disagree" to "strongly agree" (strongly

disagree = 1; disagree = 2; neutral = 3; agree= 4; strongly agree =5). The higher the score, the more anxious a student is when speaking English.

2) Reflective journal

Reflective journals, also called reflective diaries or learning journals, are valuable means for capturing participants' perspectives on the phenomena being studied [38]. Reflective journals are containers for recording thoughts, emotions, and learning progress over a specific period, helping individuals gain self-awareness and a deeper understanding of their learning journey [39].

In this study, the reflective journals offered a dynamic record of participants' anxiety reactions, perceptions, and self-assessments. Since learners may hesitate to discuss sensitive issues openly, reflective journals allow them to express these thoughts privately. The reflective journal (see Appendix Table A2) was modeled after Ahmed's design [40]. It was sent to participants weekly, inviting them to share their experiences with ASR technology and its effects on their learning process.

C. Instructional Design

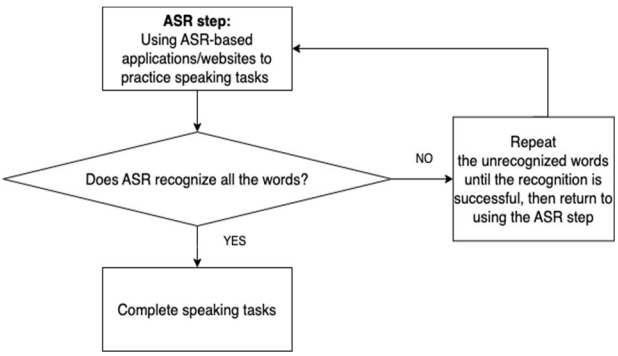

Fig. 1. Participants' workflow.

The EFL course was a compulsory foreign language course for Year 1 and Year 2 undergraduates. It focuses on mastering pronunciation, intonation, and vocabulary and improving speaking skills. The course consists of six thematic units over 14 weeks. Week 1 was an orientation session, and participants were introduced to English-speaking learning assisted by mobile ASR dictation software. Instructional practice started from week 1 to week 14. Each week, the course instructor distributes pre-class learning tasks to participants through an online learning platform called Xuexitong (https://v8.chaoxing.com/) and asks participants to master these tasks by themselves before the weekly face-to-face sessions. Pre-class learning tasks include input-oriented material (e.g., new words and expressions) and output-oriented tasks to check mastery (a pre-class speaking task related to the unit topic). They must study the content independently and finish the speaking task based on their self-study efforts to prepare. The course instructor encouraged participants to practice the speaking task repeatedly until they could deliver a three-minute monologue fluently with minimal errors (see Appendix C for an example of speaking tasks). Fig. 1 outlines the participants' process of using ASR to complete their speaking tasks. Participants used mobile ASR-enabled dictation applications to practice speaking tasks. They focused on pronouncing unrecognized words repeatedly until the ASR system accurately recognized them.

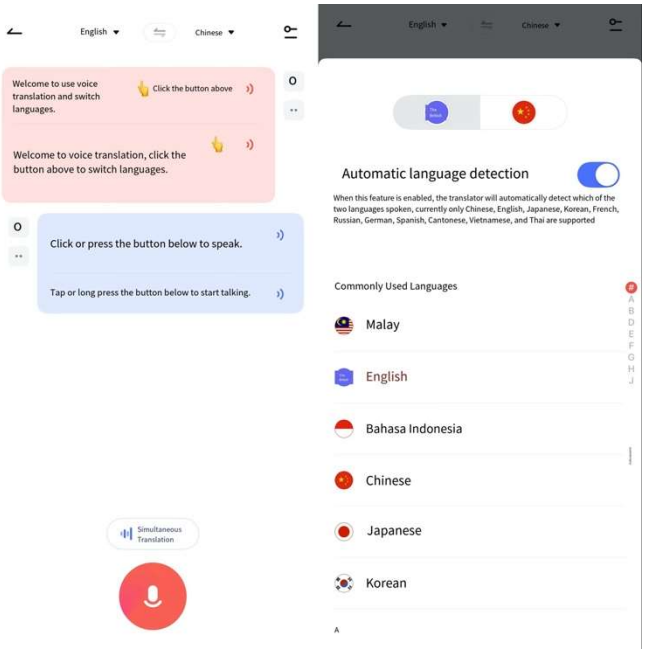

Fig. 2. Snapshot of the NetEase Youdao application interface.

Participants can choose the mobile ASR-enabled dictation applications according to their preferences or the NetEase Youdao Dictionary recommended by the course teacher to assist their English-speaking learning. NetEase Youdao Dictionary is a free dictionary application software based on mobile terminals. It works on a variety of devices, including iOS and Android. It supports the translation of 109 languages, including Chinese and English. It can transcribe and translate speech into text in various languages in real-time (see Fig. 2 for the application user interface). This app is recommended for two main reasons. First of all, the app was developed by NetEase Youdao, the first educational technology company to develop a large educational model in China and land in oral practice, and its voice transcription text function is free and unlimited for users, which is a massive advantage over other apps. Second, the app supports real-time cross-linguistic speech-to-text translation. Participants can use the spoken translation function to learn the corresponding English expressions in Chinese. This may further ease their anxiety about practicing spoken English without the help of a skilled interlocutor.

D. Data Collection Process

Participants completed the Speaking Anxiety Questionnaire at the beginning (Week 1) and the end (Week 14) of the Spring semester of 2024 using the questionnaire survey web platform Wenjuanxing (https://www.wjx.cn/). In addition, they wrote reflective journals every two weeks and submitted them to the learning platform Xuexitong.

E. Data Analysis

The data were analyzed using a series of statistical procedures. Descriptive statistics were first computed for the pre-test and post-test scores of the SAQ to summarize the distribution and variability. The normality of the data was then assessed with the Kolmogorov-Smirnov (K-S) test, which confirmed that the data followed a normal distribution. Subsequently, paired samples t-tests were conducted to examine potential differences between the pre-test and post-test scores, assessing the effect of ASR technology on

speaking anxiety. Each dimension of the SAQ was analyzed to determine whether any reductions in speaking anxiety were statistically significant. Additionally, Cohen's d was calculated to assess the intervention's effect size. All quantitative data analyses were conducted using SPSS (version 29).

In addition to the quantitative analysis, the qualitative aspect of this study involved analyzing reflective journals using thematic analysis (TA). This method identifies, analyzes, and reports patterns (themes) within qualitative data. Thematic analysis is precious for understanding participants' experiences and perceptions, as it allows for an in-depth exploration of recurring themes that emerge from the data.

The analysis followed the six phases outlined by Braun and Clarke [41]: (1) familiarizing oneself with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. Throughout this process, meticulous examination and refinement of codes facilitated the identification of broader categories and subcategories, greatly enhancing the understanding of the effects of ASR on learners' speaking anxiety. All qualitative data analyses were conducted using NVIVO (version 14).

IV. Result and Discussion

A. Findings from the Questionnaire

1) Descriptive Statistics of the Pre-test and Post-test of the SAQ

Data taken from the SAQ questionnaire during the pre-test and post-test phases was examined to determine if ASR affects Chinese EFL learners' speaking anxiety. The results are presented in the Table 3 below.

Table 3. Descriptive statistics of the pre-test and post-test of the SAQ

Stage	N	Means	Std. deviations	Std. error means
Pre-test	98	45.77	5.90	0.60
Post-test	98	43.65	5.34	0.54

Table 3 presents the descriptive statistics for the pre-test and post-test scores of the SAQ questionnaire. The mean score for the pre-test was 45.77 (SD = 5.90), while the mean score for the post-test was 43.65 (SD = 5.34). The lower mean score in the post-test suggests a reduction in speaking anxiety, indicating a potential positive impact of ASR. To quantify the effect size of this change, Cohen's d was calculated. The resulting Cohen's d value of 0.296 indicates a small effect size, suggesting that the reduction in speaking anxiety is meaningful, though not particularly strong. However, further inferential statistical analysis is required to confirm the significance of this change.

2) Paired sample statistics of the pre-test and post-test of the SAQ

Before further analysis, the normality of the data was assessed using the Kolmogorov–Smirnov (K–S) test. The results indicated that the data distribution was normal (p > 0.05). Therefore, a paired samples t-test was conducted to examine the differences in pre-test and post-test scores.

Table 4 presents significant differences in several dimensions of the SAQ questionnaire, particularly in the dimensions of Self-confidence, Making Mistakes, Teacher Input, Preparation, and Volunteering to Speak (p < 0.05), suggesting a reduction in speaking anxiety.

Table 4. Paired sample statistics of the pre-test and post-test of the SAQ

Pair	Paired test items	Mean	Std. deviations	Std. error means	Sig. (2-Tailed)
Pair 1	Pre_Q1-Post_Q1	0.18	0.90	0.08	0.03
Pair 2	Pre_Q2-Post_Q2	−0.33	1.44	0.15	0.03
Pair 3	Pre_Q3-Post_Q3	0.31	1.47	1.15	0.04
Pair 4	Pre_Q4-Post_Q4	0.24	1.11	0.11	0.04
Pair 5	Pre_Q5-Post_Q5	0.25	0.98	0.10	0.02
Pair 6	Pre_Q6-Post_Q6	−0.37	1.24	0.13	0.00
Pair 7	Pre_Q7-Post_Q7	0.25	1.03	0.10	0.02
Pair 8	Pre_Q8-Post_Q8	0.27	1.32	0.13	0.05
Pair 9	Pre_Q9-Post_Q9	0.36	1.29	0.13	0.01
Pair 10	Pre_Q10-Post_Q10	0.20	0.92	0.10	0.03
Pair 11	Pre_Q11-Post_Q11	0.27	1.90	0.11	0.02
Pair 12	Pre_Q12-Post_Q12	0.27	1.13	0.11	0.02
Pair 13	Pre_Q13-Post_Q13	−0.143	1.42	0.14	0.32
Pair 14	Pre_Q14-Post_Q14	0.38	1.25	0.13	0.00

In contrast, no significant differences were found in the dimensions of Being Singled Out to Speak (p = 0.32), Peer Pressure (p = 0.05), and Fear of Being Laughed At (p = 0.32). This suggests that these anxiety factors might not be significantly influenced by ASR technology. One possible explanation is that ASR functions as a one-way, input-only tool, where learners practice speaking to a machine rather than engaging in interactive, two-way communication. Since ASR does not offer real-time conversations or social interactions, it lacks the dynamic feedback and social context that could trigger these specific anxieties, such as the fear of being judged or laughed at by others. While ASR provides a private, non-judgmental space for practice, it may not address the social components of speaking anxiety that arise in more interactive contexts.

B. Findings from the Reflective journals

The reflective journals provided valuable insights into students' experiences with ASR in language learning. The data was analyzed using thematic analysis, which helped identify vital recurring themes in students' reflections. The main themes that emerged from the data were usage strategies and perceptions of ASR feedback, reduction in speaking anxiety, improvement in speaking competence, challenges, and suggestions for improvement.

1) Usage strategies and perceptions of ASR feedback

Students employed various strategies to integrate ASR into their language learning, such as reviewing vocabulary and preparing for presentations. For instance, student 11 mentioned, "I usually use ASR to check if I'm saying things right, especially before I have to present. It has been really useful for my assignments." Many students also mentioned repeating exercises several times to improve their

pronunciation and fluency. As student 44, "I would read aloud each sentence a few times until it sounded correct." Most students found ASR feedback beneficial, especially for refining pronunciation. Student 73 noted, "The instant feedback helped me catch mistakes right away, so I didn't keep repeating them."

2) Reduction in speaking anxiety

A common theme across the reflective journals was the reduction in speaking anxiety due to ASR's immediate, non-judgmental feedback. Students highlighted that practicing with ASR felt less stressful than speaking in front of others. Student 17 commented, "Practicing ASR is less stressful because it is objective, and I prefer to be evaluated by ASR rather than judged by my classmates or teachers." Similarly, student 36 wrote, "It feels like having a personal coach who gives me immediate feedback, allowing me to practice without feeling embarrassed." This self-paced, private practice environment seemed to help students feel more at ease and in control.

3) Improvement in speaking competence

Frequent ASR practice made students more comfortable and skilled in speaking. Many students noted a decrease in hesitation and an increase in overall confidence when speaking. Student 51 said, "I've become much less nervous when speaking… Before, I would hesitate, but now it feels easier to say what I want to say." Student 73 mentioned feeling more comfortable in challenging situations, explaining, "I prefer to speak in front of others now." These reflections suggest that regular ASR practice helped students become more confident speakers.

4) Challenges

Although most experiences were positive, some students encountered challenges, such as ASR misinterpreting certain words. Student 32 wrote, "Sometimes the app doesn't understand my pronunciation, especially with words that are a bit harder for me." Student 39 shared that focusing too much on pronunciation sometimes threw off their rhythm, "I'd get caught up trying to say certain words perfectly, and it would make my speaking feel less natural." These challenges suggest that while ASR technology can be helpful, technical limitations and students' tendency to focus on perfection may affect the overall fluency of their speech.

5) Suggestions for improvement

Several students suggested ways to improve ASR, including more personalized feedback and the ability to track recurring mistakes. Student 20 mentioned, "I wish the app could keep track of the mistakes I make a lot and help me focus on fixing those." Others recommended adding more variety to the practice tasks to make them more like real conversations. Student 61 stated, "It'd be more realistic if there were different kinds of scenarios to practice, not just single phrases or sentences." Student 95 shared, "It would be helpful to have a feature that shows my progress, like how much I have improved over time. That way, I could see how far I've come.

C. Discussion

1) Addressing the Research Question 1 and Hypothesis

Research Question 1: To what extent does ASR technology reduce speaking anxiety among Chinese English majors?

Hypothesis: ASR technology significantly reduces the overall speaking anxiety of Chinese English majors, as measured by the SAQ questionnaire.

The paired samples t-test results show a statistically significant reduction in anxiety scores from the pre-test to the post-test, indicating a general decrease in speaking anxiety among participants. The findings support this study's hypothesis—that ASR technology significantly reduces the overall speaking anxiety of Chinese English majors. It demonstrates that ASR technology contributes to overall reduced speaking anxiety in this context. This reduction can likely be attributed to the immediate, objective feedback provided by ASR, which allows students to practice speaking in a private, non-judgmental setting, reducing the anxiety associated with making mistakes in front of others.

2) Addressing the Research Question 2

Research Question 2: How do students perceive and experience ASR as a tool for reducing speaking anxiety?

Reflective journals provide additional insights into students' experiences with ASR. Many students reported feeling more comfortable and less anxious due to ASR's non-judgmental, private nature. Students described ASR as "a personal coach" for practice without fear of embarrassment, indicating that ASR's design fosters confidence and reduces stress in language practice.

3) Integration of Quantitative and Qualitative Findings

The quantitative results from the SAQ and the reflective journals confirm that ASR is an effective tool for reducing speaking anxiety and addressing language learners' psychological and practical needs. The integration of these two approaches highlights that ASR not only reduces anxiety but also helps learners become more confident in their speaking abilities, providing both a tangible reduction in anxiety scores and subjective reports of improved self-confidence. These findings suggest that ASR could be a valuable resource in EFL settings, supporting speaking skill improvement and a more comfortable, anxiety-reducing learning environment.

4) Limitations

While this study provides valuable insights into using ASR technology to reduce speaking anxiety among Chinese EFL learners, several limitations exist. First, the sample was limited to a specific group of Chinese English majors, which may restrict the generalizability of the findings to other student populations or cultural contexts. Second, the study relied on self-reported data through reflective journals, which may introduce subjective bias. Additionally, the study only measured short-term effects; therefore, it is still being determined if the anxiety-reducing effects of ASR are sustained over time. Future research could address these limitations by including a more diverse sample, using a longitudinal design, and incorporating objective measures of speaking performance.

V. Conclusion

This study proves that ASR technology effectively reduces speaking anxiety among Chinese English majors, affirming the initial hypothesis. The statistically significant decrease in SAQ scores from the pre-test to the post-test reflects a tangible reduction in anxiety. This suggests that ASR can

alleviate students' psychological barriers in speaking practice.

The qualitative insights from students' reflective journals further underscore ASR's role in creating a supportive and private practice environment. Students valued the chance to practice without the fear of judgment, describing ASR as a "personal coach" that allows for self-paced learning. This unique characteristic of ASR technology appears crucial in fostering confidence and skill improvement, as students felt less embarrassed and more willing to engage in speaking activities.

The integration of quantitative and qualitative findings points to ASR as a potentially valuable resource in EFL settings, particularly where speaking anxiety is prevalent. Addressing psychological and practical learning needs, ASR helps create a more effective and anxiety-reducing environment for language practice.

Future research should address these limitations by expanding to diverse populations, exploring long-term impacts, and integrating objective performance measures. Further investigations could also explore ways to optimize the use of ASR for anxiety reduction, such as through personalized feedback and more varied practice scenarios, enhancing its potential for improving speaking competence.

In conclusion, this study demonstrates that ASR technology offers a promising approach to reducing speaking anxiety and improving speaking competence among EFL learners. The findings suggest that ASR can be a practical addition to language learning environments, supporting students in building both skill and confidence as they navigate the challenges of speaking a second language.

APPENDIX

Table A1. Speaking anxiety questionnaire

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	2	3	4	5

No.	Items
1	I never feel quite sure of myself when I am speaking English in my class
2	I don't worry about making mistakes in the English class.
3	It frightens me when I don't understand what the teacher is saying in English.
4	I start to panic when I have to speak without preparation in English class.
5	It embarrasses me to volunteer answers in my English class.
6	I feel confident when I speak English in class.
7	I am afraid that my English teacher is ready to correct every mistake I make.
8	I can feel my heart pounding when I am going to be called on in the English class.
9	I always feel that the other students speak English better than I do.
10	I feel very self-conscious about speaking English in front of other students.
11	I get nervous and confused when I am speaking English in class.
12	I get nervous when I don't understand every word the English teacher says.
13	I am afraid that the other students will laugh at me when I speak English.
14	I get nervous when the English teacher asks questions which I haven't prepared in advance.

Table A2. Reflective journal

No.	Items
1	How do you learn or prepare yourself before coming to the classroom?
2	What do you think of the feedback provided by ASR?
3	How do you feel about the implementation of ASR? How have your feelings changed?
4	What are the strengths and weaknesses that you can identify after three weeks of the implementation of ASR?
5	What are the challenges that you have faced during the implementation of ASR? How did you deal with it?
6	Please evaluate the impact of ASR on your speaking anxiety.
7	After using ASR, how do you feel about your English speaking anxiety?
8	Will you consider using ASR to assist your language learning in the future? Why?
9	To what do you recommend using ASR for the next semesters? Why?

Unit 3: Inquiry and Offer

Pre-class ASR-Assisted Oral Task:

Imagine you are interested in purchasing a new product or service, such as a computer, or a membership. Prepare a two-minute conversation as if you are calling or speaking to a representative to ask for more information. Use the following questions to help guide your inquiry:

- What product or service are you interested in?
- What specific information do you need (e.g., features, price, availability, delivery options)?
- What additional questions would you ask to confirm or clarify details?
- How would you respond if the product or offer does not meet your expectations?

In-Class Unit Task:

Scenario: Imagine you need to purchase a new electronic device for school or work. You visit a store or call a company to inquire about the options available within your budget and specific needs.

Roles and Setup:

Student A (Customer): You have a specific budget and list of desired features. Politely inquire about prices, features, availability, and any potential discounts or offers. Clearly express your needs and preferences and be ready to consider alternatives if your preferred option isn't available.

Student B (Sales Representative): Respond to the customer's inquiries by offering several options within their budget and needs. Highlight special features, compare options, and, if needed, suggest alternatives. You might also provide a discount or special deal if appropriate.

Student C (Customer's Friend): Help the customer make an informed decision by asking additional questions and providing advice based on the evaluation criteria (e.g., cost, quality, customer reviews, and usability).

Follow-up Task:

After the role-play, each group should discuss which product or service they chose and why. Then, write a short email thanking the representative and confirming your final choice, or ask further questions if you are undecided.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

W.Y.L., M.B.M., and H.W.Y conducted the research, analyzed the data, and wrote the paper. All authors had approved the final version.

REFERENCES

[1] H. Xu, "Optimization of English online teaching scheme based on cloud platform technology," *Mathematical Problems in Engineering,* vol. 2022, no. 1, 2493632, 2022.

[2] A. Gupta, R. Kumar, and Y. Kumar, "Hybrid deep learning based automatic speech recognition model for recognizing non-Indian languages," *Multimedia Tools and Applications,* vol. 83, no. 10, pp. 30145–30166, 2024.
[3] K.-W. K. Lai and H.-J. H. Chen, "An exploratory study on the accuracy of three speech recognition software programs for young Taiwanese EFL learners," *Interactive Learning Environments,* vol. 32, no. 5, pp. 1582–1596, 2024.
[4] H. Xiao, K. Ou, H. Wang, and J. Weijer, "The effect of ASR apps on monophthong pronunciation improvement and generalization to new words in English," in *Proc. Science and Information Conference*, 2023, Springer, pp. 1410–1433.
[5] K. Evers and S. Chen, "Effects of automatic speech recognition software on pronunciation for adults with different learning styles," *Journal of Educational Computing Research,* vol. 59, no. 4, pp. 669–685, 2021.
[6] R. P. Yaniafari and V. Olivia, "The potential of ASR for improving English pronunciation: A review," *KnE Social Sciences,* pp. 281–289-281–289, 2022.
[7] M. Y.-C. Jiang, M. S.-Y. Jong, W. W.-F. Lau, C.-S. Chai, and N. Wu, "Effects of automatic speech recognition technology on EFL learners' willingness to communicate and interactional features," *Educational Technology & Society,* vol. 26, no. 3, pp. 37–52, 2023.
[8] S. M. McCrocklin, "Pronunciation learner autonomy: The potential of automatic speech recognition," *System,* vol. 57, pp. 25–42, 2016.
[9] M. Bashori, R. van Hout, H. Strik, and C. Cucchiarini, "Web-based language learning and speaking anxiety," *Computer Assisted Language Learning,* vol. 35, no. 5–6, pp. 1058–1089, 2022.
[10] E. K. Horwitz, M. B. Horwitz, and J. Cope, "Foreign language classroom anxiety," *The Modern language journal,* vol. 70, no. 2, pp. 125–132, 1986.
[11] I. M. Pabro-Maquidato, "The experience of English speaking anxiety and coping strategies: A transcendental phenomenological study," *International Journal of TESOL & Education,* vol. 1, no. 2, pp. 45–64, 2021.
[12] S. Jin, "Speaking proficiency and affective effects in EFL: Vlogging as a social media-integrated activity," *British Journal of Educational Technology,* vol. 55, no. 2, pp. 586–604, 2024.
[13] E. Sudina, "Scale quality in second-language anxiety and WTC: A methodological synthesis," *Studies in Second Language Acquisition,* vol. 45, no. 5, pp. 1427–1455, 2023.
[14] K. Piniel, "Language anxiety," *Investigating Foreign Language Anxiety: Lessons for Research into Individual Differences*, Springer, 2024, pp. 7–18.
[15] X. Zhao, "A review on the effects of foreign language anxiety on second language learning," *International Journal of Linguistics, Literature and Translation,* vol. 6, no. 2, pp. 150–156, 2023.
[16] Z. Zhong, "A study on the current situation of English learning anxiety of English majors," *Journal of Education, Humanities and Social Sciences,* vol. 23, pp. 44–50, 2023.
[17] W. Sun, "The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation," *Frontiers in Psychology,* vol. 14, 1210187, 2023.
[18] K. Evers and S. Chen, "Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults," *Computer Assisted Language Learning,* vol. 35, no. 8, pp. 1869–1889, 2022.
[19] S.-C. Tsai, "Learning with mobile augmented reality-and automatic speech recognition-based materials for English listening and speaking skills: Effectiveness and perceptions of non-English major English as a foreign language students," *Journal of Educational Computing Research,* vol. 61, no. 2, pp. 444–465, 2023.
[20] Y. E. Yesilyurt, "AI-enabled assessment and feedback mechanisms for language learning: Transforming pedagogy and learner experience," *Transforming the Language Teaching Experience in the Age of AI*, IGI Global, 2023, pp. 25–43.
[21] M. Y.-C. Jiang, M. S.-Y. Jong, W. W.-F. Lau, C.-S. Chai, and N. Wu, "Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom," *Australasian journal of educational technology,* vol. 37, no. 2, pp. 110–131, 2021.
[22] D. S. Dhivya, A. Hariharasudan, W. Ragmoun, and A. A. Alfalih, "ELSA as an education 4.0 tool for learning business English communication," *Sustainability,* vol. 15, no. 4, p. 3809, 2023.
[23] A. Kholis, "Elsa speak app: automatic speech recognition (ASR) for supplementing English pronunciation skills," *Pedagogy: Journal of English Language Teaching,* vol. 9, no. 1, pp. 01–14, 2021.
[24] Y. Hayashi, Y. Kondo, and Y. Ishii, "Automated speech scoring of dialogue response by Japanese learners of English as a foreign language," *Innovation in Language Learning and Teaching,* vol. 18, no. 1, pp. 32–46, 2024.
[25] S. McCrocklin, C. Fettig, and S. Markus, "Salukispeech: Integrating a new ASR tool into students' English pronunciation practice," *Pronunciation in Second Language Learning and Teaching Proceedings,* vol. 12, no. 1, 2022.
[26] E. K. Horwitz, M. Tallon, and H. Luo, "Foreign language anxiety," *Anxiety in Schools: The Causes, Consequences, and Solutions for Academic Anxieties,* vol. 2, pp. 96–115, 2010.
[27] K. J. Hall, K. Ooteghem, and W. E. McIlroy, "Emotional state as a modulator of autonomic and somatic nervous system activity in postural control: A review," *Frontiers in neurology,* vol. 14, 1188799, 2023.
[28] K. Sho, "Comparison of a videoconferencing intervention's effects on students' English-speaking anxiety," in *Proc. 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2020, IEEE, pp. 530–535.
[29] H. N. Nguyen, "Mobile phones' video recording tool: A solution to freshmen's English-speaking anxiety," *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT),* vol. 11, no. 2, pp. 16–32, 2021.
[30] A. Rahman and P. Tomy, "Intelligent personal assistant—An interlocutor to mollify foreign language speaking anxiety," *Interactive Learning Environments,* pp. 1–18, 2023.
[31] U. Nasri, "Exploring qualitative research: A comprehensive guide to case study methodology," *Al-Hikmah: Jurnal Studi Islam,* vol. 4, no. 3, pp. 72–85, 2023.
[32] K. Dehalwar and S. N. Sharma, "Exploring the Distinctions between Quantitative and Qualitative Research Methods," *Think India Journal,* vol. 27, no. 1, pp. 7–15, 2024.
[33] M. Liu, "Understanding Chinese middle school students' anxiety in English speaking class," *Journal of Asia TEFL,* vol. 15, no. 3, p. 721, 2018.
[34] E. K. Horwitz, "On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners," *New Insights into Language Anxiety: Theory, Research and Educational Implicaations,* vol. 31, p. 47, 2017.
[35] E. K. Horwitz, "Factor structure of the foreign language classroom anxiety scale: Comment on Park (2014)," *Psychological Reports,* vol. 119, no. 1, pp. 71–76, 2016.
[36] M. Liu, "Anxiety in Chinese EFL students at different proficiency levels," *System,* vol. 34, no. 3, pp. 301–316, 2006.
[37] J.-M. Dewaele and P. D. MacIntyre, "The two faces of Janus? Anxiety and enjoyment in the foreign language classroom," *Studies in Second Language Learning and Teaching,* vol. 4, no. 2, pp. 237–274, 2014.
[38] P. Leone, S. Aranha, and S. Marques Spatti Cavalari, "Our interaction was very productive": Levels of reflection in learners' diaries in teletandem, *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication,* vol. 26, no. 3, 2023.
[39] D. D. Stevens and J. E. Cooper, *Journal Keeping: How to Use Reflective Writing for Learning, Teaching, Professional Insight and Positive Change*, Taylor & Francis, 2023.
[40] A. M. Ahmed, "From reluctance to addiction: The impact of reflective journals on Qatari undergraduate students' learning," *Reflective Practice,* vol. 21, no. 2, pp. 251–270, 2020.
[41] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology,* vol. 3, no. 2, pp. 77–101, 2006.