International Journal of Learning Technologies and Learning Environments International
International Institute of Applied Informatics
2025, Vol. 8, No. 1, IJLTLE905

Investigating the Impact of an English Meeting Transcription Tool on Student Speaking Skills and Confidence in EMI Contexts

Wen-Chun Lin [*], Yu-Chi Yeh [†], Chih-Kai Chang [*]

Abstract

With globalization, English proficiency is crucial. Taiwan widely adopts English Medium Instruction (EMI) to prepare students for the global job market. However, many learners struggle with spoken English. This study utilizes an English meeting transcription tool to help students identify and improve their speaking skills, such as pronunciation, grammar, and fluency. By providing visual feedback, the tool encourages self-directed learning and reduces speaking anxiety. The study integrated this tool with the jigsaw cooperative learning method, promoting cooperative learning and active participation. Results from eight students showed improved idea expression and reduced speaking anxiety, demonstrating the effectiveness of this approach. The study also explored students' perceptions through questionnaires and interviews, focusing on their speaking confidence, past experiences, and the impact of the assistive tool. Students found the tool user-friendly and the activities engaging, leading to increased participation and motivation. The findings suggest that technology-enhanced learning environments positively impact speaking confidence and recommend the use of transcription tools in English courses.

Keywords: English Medium Instruction, transcription tool, speaking confidence, technology-enhanced language learning.

1 Introduction

English learning typically revolves around four core skills: listening, speaking, reading, and writing. However, Taiwan's educational system has traditionally placed greater emphasis on listening and reading, leaving speaking and writing skills underdeveloped. This long-standing imbalance has resulted in a significant challenge: while many Taiwanese students can understand English through reading and listening, they often lack the confidence to use the language in real-life conversations. The lack of speaking practice creates a substantial barrier, as students frequently find themselves struggling to articulate their thoughts clearly in unfamiliar or pressured situations. This issue is further complicated by the influence of students' mother tongue, which introduces

[*] National University of Tainan, Tainan, Taiwan
[†] National Cheng Kung University, Tainan, Taiwan

challenges in pronunciation, grammar, and syntax, affecting both fluency and accuracy in spoken English. The problem is exacerbated by Taiwan's exam-oriented education system, where students focus primarily on test performance rather than practical language application. As a result, students are seldom exposed to authentic speaking opportunities that build their confidence in conversational English. Research from Indonesia highlights similar challenges among EFL (English as a Foreign Language) learners, whose pronunciation is heavily influenced by their native language. This linguistic interference makes it difficult for them to express themselves naturally in English [15]. A comparable phenomenon is prevalent in Taiwan, where many students, lacking confidence in their pronunciation, tend to avoid speaking English altogether. Pronunciation is a critical component of effective English communication, as errors in pronunciation can lead to misunderstandings, create negative impressions, and hinder communication efficiency.

To address these issues, this study aims to develop innovative English-speaking tools and teaching activities that transcend the limitations of traditional learning methods. These tools are designed to provide learners with the flexibility to practice speaking English anytime and anywhere, effectively removing barriers of time and location. By integrating modern technologies, the approach significantly increases students' opportunities for speaking practice, helping them build confidence and improve fluency over time. The study emphasizes the importance of consistent practice and a supportive environment in overcoming pronunciation challenges. By creating a more interactive and engaging learning experience, these tools and methods enable learners to develop their speaking skills more naturally and effectively. Ultimately, the goal is to help students achieve clear, confident communication in English, equipping them for both academic and professional success in a globalized world.

Technology has become indispensable in language learning, offering a dynamic and efficient approach that complements traditional methods. By providing immediate feedback and enabling repetitive practice, technology fosters a closed-loop learning process of "error-feedback-improvement," significantly enhancing learning efficiency. Moreover, technology transcends the limitations of time and space, allowing learners to access language learning tools anytime and anywhere. This flexibility accommodates diverse learning styles and paces. Through interactive elements such as graphs and maps, learners can visually track their progress and identify areas for improvement. Additionally, speech recognition tools enable real-time comparison of learners' pronunciation with native accents, contributing to enhanced pronunciation accuracy and overall language proficiency.

Technology-assisted learning showcases its unique value in language acquisition by addressing the limitations of traditional teaching methods while offering learners more flexible and efficient pathways. Learners can identify issues in real time and repeatedly practice to form a "mistake-feedback-improvement" loop, which significantly enhances their learning efficiency. Moreover, technology eliminates the time and location constraints of traditional classrooms, enabling learners to practice anytime and anywhere. This flexibility meets the diverse needs of learners with varying schedules and learning preferences. Visual tools, such as charts and interactive maps, present the learning process in a clear and intuitive way, helping learners track their progress and identify areas for improvement. For instance, pronunciation tools can convert learners' speech into visual representations, compare it with standard American pronunciation, and provide immediate feedback. This approach helps learners refine their pronunciation accuracy and enhance their overall language expression skills. By combining real-time feedback with flexible learning opportunities, technology-assisted learning creates an engaging and effective

environment for language acquisition, empowering learners to achieve their goals with confidence and precision.

Building on the research motivation outlined earlier, this study explores the use of speech recognition software as a tool to support English-speaking practice. The primary goal is to improve students' pronunciation accuracy while boosting their confidence in speaking English. This is achieved through cooperative learning, a strategy where students work in small groups to support and learn from each other. To guide the investigation, the study has identified three key research questions.

1. How does the identification of pronunciation mistakes by speech recognition software influence students' confidence in speaking?
2. How can speech recognition in cooperative learning be used to empower students with the self-confidence to speak boldly in front of their peers?
3. To what extent does the use of speech recognition software in speaking practice affect students' confidence compared to traditional teaching methods?

This study aims to explore the impact of speech recognition software on improving students' speaking skills. By designing effective teaching methods, it seeks to understand how this technology enhances pronunciation, increases speaking practice opportunities, and ultimately boosts students' confidence in speaking.

2 Literature Review

2.1 Cooperative Learning and Confidence

Cooperative learning is essential in language acquisition, addressing the limitations of traditional, teacher-centered methods. By fostering a cooperative environment, students become active participants rather than passive recipients of information, enhancing engagement and motivation. Group tasks, tailored to individual strengths and interests, promote interdependence and teamwork. This cooperative process not only boosts confidence but also provides ample opportunities for authentic language use, mitigating anxiety often associated with language learning. Unlike solitary study, cooperative learning creates a supportive peer community where learners can practice language skills in a low-stakes environment, gradually overcoming language barriers and developing fluency.

Cooperative learning emphasizes that students achieve learning goals through mutual cooperation during the teaching process. First introduced in the late 1800s, this model views students as active participants, with the teacher acting as a guide. Through team discussions and cooperative activities, students help each other acquire knowledge [16]. Cooperative learning can be applied across various subjects, including scientific exploration, reading and writing, mathematics, and science [8][4]. Additionally, experimental studies on cooperative learning suggest that learning through cooperation can effectively increase students' motivation to learn. Scholars have proposed various strategies for implementing cooperative learning. Below are brief descriptions of three well-known strategies.

Three collaborative learning methods are described. First, the Jigsaw method, developed by Slavin, divides students into diverse groups, assigns individual research topics, and facilitates expert group discussions before returning to the original group for collective learning [7]. Second, Cognitive Apprenticeship, proposed by Brown, Collins, and Duguid, emphasizes problem-solving through structured activities, fostering knowledge application and framework development [5]. Lastly, Teams-Games-Tournament (TGT) involves heterogeneous group practice followed by quiz competitions with rewards based on rankings. In the described experiment, the Jigsaw method was integrated to enhance learning, aiming to increase student motivation and participation through teamwork and support.

Cooperative learning strategies, such as jigsaw, cognitive apprenticeship, and team-games-tournament, foster a cooperative environment where students actively engage in learning and support each other. By working together in heterogeneous groups, students have ample opportunities to practice their language skills, build confidence, and overcome any language barriers they may have. For instance, in a language learning context, the cognitive apprenticeship approach can be adapted to simulate real-life conversations, allowing students to practice their speaking skills in a low-stakes environment. Through these interactive learning experiences, students develop the necessary skills to express themselves confidently and effectively in a foreign language. Moreover, the social interaction inherent in cooperative learning helps students to overcome their fear of making mistakes and build their self-esteem. As they work together to achieve common goals, students develop a sense of belonging and camaraderie, which can significantly boost their confidence in speaking in front of others.

Confidence refers to an individual's self-assessment of their ability to achieve a specific task or goal, reflecting a positive and proactive psychological state. Confidence influences persistence and future success [12]. In Keller's ARCS motivation model proposed in 1983, confidence is a key component [11]. Keller (1987) outlined five strategies to enhance confidence, including:

1. Learning Requirements: Clearly define learning objectives and ensure they are supported by appropriate assessment methods.
2. Difficulty: Provide challenges that are suitably demanding, encouraging students to push their limits.
3. Expectations: Assist students in developing practical and attainable learning plans.
4. Attributions: Help students recognize the factors contributing to both their successes and setbacks.
5. **Self-Confidence**: Enable students to independently adjust their mindset in pursuit of their learning goals. *Collaborative learning boosts confidence by providing peer support and shared success*. Explaining concepts reinforces understanding, while positive feedback from peers validates individual contributions, reducing anxiety and fostering a sense of competence.

Linguistic self-confidence, or L2 confidence, refers to an individual's self-assessment of their language abilities and a lack of anxiety [6]. A study by Apple (2011) suggested that learners' confidence in speaking English is influenced by their perception of a positive classroom experience [2]. Since confidence is an internal feeling that is difficult to observe directly, reliable and valid questionnaires are needed to track its changes. In Apple's study, a 30-item Foreign Language Classroom Speaking Confidence (FLCSC) questionnaire was introduced to assess students' confidence in speaking English and the factors that influence it.

2.2 Transcription Tools for EFL Speaking

Phonetic transcription is a valuable tool in language learning, particularly for improving pronunciation skills. It has been shown to significantly enhance students' pronunciation, especially in terms of segmental features such as consonants and vowels, and suprasegmental features like word stress and sentence stress. This is because phonetic transcription provides a visual representation of speech sounds, allowing learners to see how words are pronounced and identify their own pronunciation errors. Studies have demonstrated that using phonetic transcription as a teaching method can lead to significant improvements in pronunciation, particularly among students in experimental groups. For instance, the study conducted at Pamulang University in the academic year 2019/2020 found that fourth-semester English Literature students showed marked improvement in their pronunciation after being taught using phonetic transcription [10]. While phonetic transcription may have limitations in improving intonation, it can still be an effective tool for enhancing pronunciation when combined with other teaching strategies. By using phonetic transcription, students can gain a deeper understanding of pronunciation, build confidence in their speaking abilities, and achieve more fluent and natural English speech.

Language learners can enhance their pronunciation skills through the phonetic transcription of television news, which serves as an engaging and effective method for practice. Listening to news broadcasts is not only stimulating for learners but also allows them to record and replay segments as needed, facilitating repeated exposure to accurate pronunciation. By transcribing the spoken content, learners can focus on mastering the correct articulation of words. The use of speech transcription tools significantly aids students in grasping word stress and sentence intonation, particularly in the context of connected speech, thereby improving fluency and naturalness in their spoken language. These skills form a crucial foundation for effective communication in language learning. Research indicates that students can regularly record their speaking practice and compare multiple transcriptions to track their progress, which in turn enhances their motivation to learn [3].

The study conducted by Maulia Hafizhah and colleagues emphasizes the significant role of transcription in enhancing English pronunciation through technological assistance, particularly via the ELSA Speaking Application [9]. The research highlights that phonetic transcription serves as a critical tool in language learning, enabling students to grasp the nuances of English pronunciation more effectively. By providing accurate phonetic representations, the ELSA application facilitates learners' understanding of sound production, which is essential for clear communication. Furthermore, the integration of technology in language education not only makes learning more engaging but also allows for personalized feedback through automatic speech recognition features. This combination of transcription and technology not only aids in overcoming common pronunciation challenges faced by learners but also fosters a more conducive learning environment, ultimately enhancing students' confidence and proficiency in spoken English. The findings indicate that students perceive the ELSA Speaking Application as an effective resource for mastering both pronunciation and phonetic transcription, thereby underscoring the broader implications of technology-assisted language learning in modern education.

2.3 Examples of English-speaking Related Learning Systems- ELSA Speak

ELSA Speak is an English-speaking learning application developed by ELSA Corp. in San Francisco, USA, in 2015. The app breaks pronunciation learning into small units, enabling learners to practice from words and phrases to short sentences and conversations. Its interface includes features for learning vowels, consonants, and stressed syllables (see Figure 1, left). ELSA Speak's strength lies in its color-coded feedback system. When learners record their voice, the app uses speech recognition technology to compare it with standard American English pronunciation. Syllables are color-coded green, yellow, or red, indicating pronunciation accuracy as correct, close, or off (see Figure 1). If learners are unsure about a specific word's pronunciation, they can click on it for a detailed syllable-by-syllable explanation (see Figure 1). This feedback system helps learners identify correct pronunciation areas and areas needing improvement, enhancing learning through repeated practice.

ELSA Speak was used in the English education program at STKIP Muhammadiyah Enrekang University in Indonesia during 2018 and 2019. The study, which involved 12 students in a pronunciation course focused on specific themes, revealed positive feedback on the system's content design, assessment methods, interactivity, and speech recognition. Students appreciated ELSA Speak's design and expressed interest in using it in future speaking courses [14]. Based on this research, it is clear that the color-coded feedback system in English-speaking tools effectively boosts students' learning motivation. Consequently, the English-speaking assistance system developed in this research also incorporated these features.

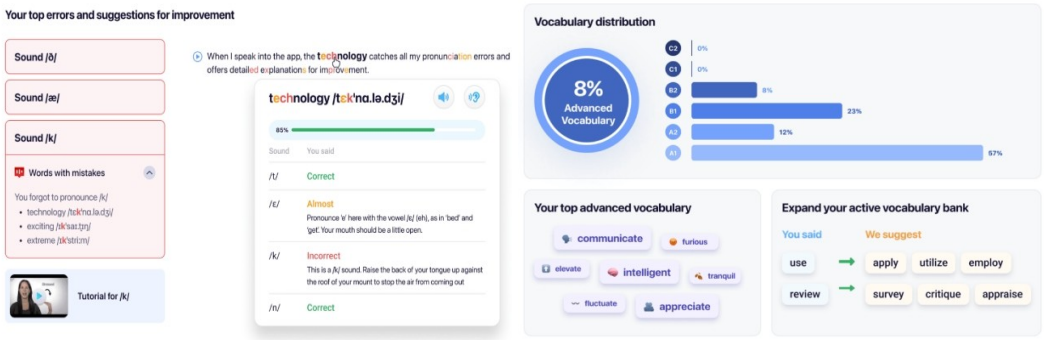

Figure 1: Instantly correct pronunciation errors (left) and vocabulary (right).

3 Research Methodology and Tools

3.1 Research Tool - Otter.ai

Unlike traditional speech recognition tools, Otter.ai (formerly AISense) is an AI-powered transcription tool that converts audio into highly accurate and searchable text. Unlike basic speech-to-text conversion, Otter.ai utilizes advanced verbatim transcription technology to deliver precise text records. Additionally, its built-in speaker identification and keyword search features make it particularly useful for language learning, allowing students to review and refine their pronunciation—an advantage that traditional speech recognition tools often lack.. This feature makes it an

invaluable resource for language learners seeking to improve their spoken English. By transcribing their own speech or listening to English audio, learners can gain valuable insights into their pronunciation, grammar, and fluency. Otter.ai's ability to generate detailed transcripts allows users to easily identify areas for improvement and focus on specific language features. Moreover, the platform's integration with popular video conferencing tools like Zoom enables learners to transcribe their online language exchange sessions or lessons, providing a comprehensive record of their progress. With Otter.ai, language learners can effectively track their development, identify patterns in their errors, and ultimately enhance their spoken English skills.

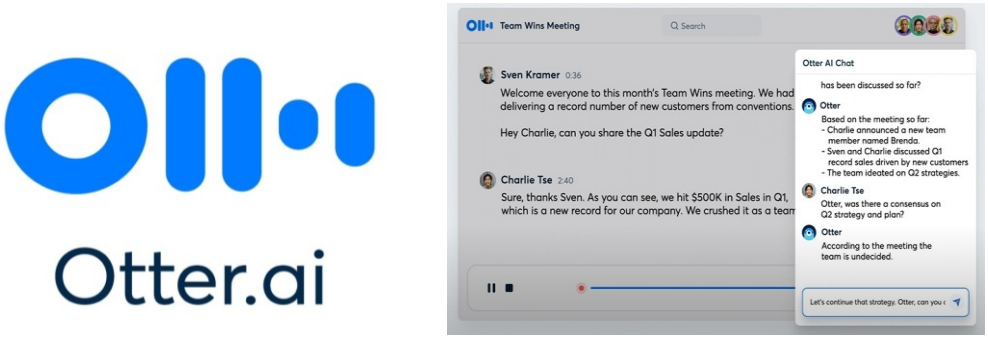

Figure 2: Accurate recording pronunciation errors (left) and vocabulary (right).

3.1.1 Why Otter.ai?

This study selects Otter.ai for three main reasons. First, it provides complete verbatim transcription rather than merely recognizing individual words, allowing learners to fully review their spoken content. Second, it supports speaker identification, making it suitable for collaborative learning models and helping students reflect on their pronunciation and expression during group discussions. Third, it offers cloud access, enabling students to review their learning progress anytime, thereby enhancing the potential for autonomous learning (see Table 1).

Otter.ai is a platform that records audio and transcribes it into text in real-time, as shown in Figure 2 (right). Using artificial intelligence and speech recognition technology, Otter.ai automatically transcribes spoken content, making it easy to search, review, and share meeting details. It facilitates online collaboration and learning by rapidly converting audio into text. Supporting multiple languages, including English, Mandarin, Cantonese, Japanese, and Korean, Otter.ai can recognize different speakers and background noise. Otter.ai is versatile and can be used for note-taking, organizing meeting records, and creating textual content. Its features, such as keyword search, audio recording, and annotation, enhance productivity and accuracy. By reducing the time spent searching for previous meeting content and textual records, Otter.ai saves time and costs. It is suitable for various scenarios, such as business meetings, online learning, language practice, speeches, and interviews. Otter.ai's automatic transcription feature improves efficiency and reduces the burden of manual recording, ensuring that users can keep up with discussions. Additionally, Otter.ai is available across multiple platforms, including computers, phones, and tablets (iOS, Android, and web), allowing convenient access anytime, anywhere.

Feature	Otter.ai	Google Speech-to-Text	ELSA Speak	Speechling
Real-time verbatim transcription	☑	✗	✗	✗
Speaker identification	☑	✗	✗	✗
Keyword search	☑	☑	✗	✗
Zoom integration	☑	✗	✗	✗
Pronunciation scoring	✗	✗	☑	☑
Suitable for group learning	☑	✗	✗	✗

Table 1: Comparison of Speech Transcription and Pronunciation Assessment Tools

3.2 Research Tool - FLCSC Questionnaire

In 2011, Apple introduced a 30-item Foreign Language Classroom Speaking Confidence (FLCSC) questionnaire to assess students' confidence in speaking English in the classroom. The questionnaire also examined the relationships between the Big Five Personality Traits, students' English-speaking confidence, and the social environment. It consisted of six dimensions:

1. Foreign Language Classroom Speaking Anxiety (FLCSA)
2. Perceived Foreign Language Speaking Self-Competence (PFLSS)
3. Desire to Speak English (DSE)
4. Past English Classroom Experiences (PECE)
5. Current English Classroom Perception (CECP)
6. Perceived Social Value of Speaking English (PSV)

In Apple's (2011) study, the PECE (Social Support in Educational Environment) and PSV (Perceived Value of Language Learning) dimensions of the questionnaire used a 4-point Likert scale, while the remaining dimensions used a 6-point Likert scale. Based on the questionnaire results, the first three dimensions (FLCSA, PFLSS, DSE) represented speaking confidence, while the last three dimensions (PECE, CECP, PSV) were identified as factors influencing speaking confidence. The reliability of the scales was high, with in-class anxiety at .89, out-of-class anxiety at .87, and the combined scales at .94, indicating that the FLCSA is a reliable tool. Apple (2011) recommended increasing communication opportunities in the classroom, such as pair or group work, to provide students with more speaking practice. Teachers should also offer greater encouragement and guidance to help boost students' speaking confidence and reduce speaking anxiety.

3.4 Research Design

As shown in Figure 3, this study involved a loop structure with five steps of student English-speaking presentations. Each week, students would be assigned content from the learning text and would take turns presenting every other week. The study incorporated the jigsaw method from cooperative learning, combined with the English meeting transcription tool for classroom activities. The jigsaw method's rules and reward mechanism were applied during this process. Using the transcription tool, pronunciation errors can be identified, allowing for individual speaking assessments. Students can compare the recognized text with the actual textbook text to reflect on their pronunciation. Through repeated practice, students can improve the accuracy of their English pronunciation. In this cooperative learning environment, students would gradually become more willing to practice speaking and communicate with others, building confidence in learning English. After using the transcription tool, the study would conduct one-on-one qualitative interviews based on the Foreign Language Classroom Self-Confidence Scale (FLCSC) questionnaire. Data for the experimental psychology research were collected using PsychoPy [13]. The primary aim is to explore changes in participants' English-speaking confidence after using the transcription tool and to identify the factors influencing their confidence.

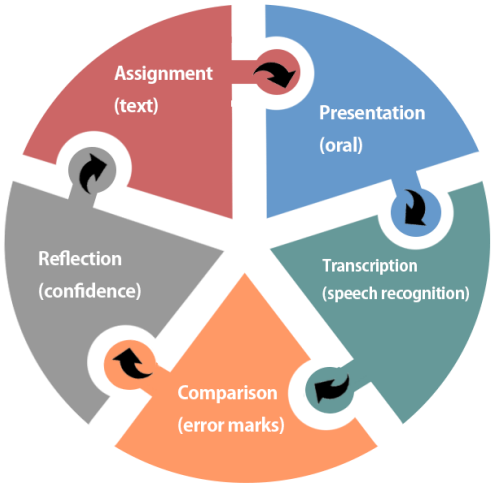

Figure 3: Loop structure of English presentation studies each week.

3.5 Learning Activity Flow

Figure 4 illustrates the activity flow, which integrates the jigsaw method of cooperative learning with the Otter.ai meeting transcription tool to enhance students' speaking confidence in a cooperative, technology-supported environment. Before class, the teacher prepares the lesson, including the learning topic, activity materials, and discussion questions to encourage student participation. The activity materials consist of a 3–5-minute speaking video and the text for student presentations. The classroom adopts a small group game competition method from cooperative learning, with three groups of 3 persons, including students and the teacher. At the start of the class, the teacher delivers a lecture, introducing the main content and key points from the video, followed by an explanation of the jigsaw method rules. The teacher

then plays the instructional video. Afterward, students are divided into groups to familiarize themselves with using the Otter.ai meeting transcription tool. Once students are comfortable with the tool, they would participate in speaking presentations. During presentations, students press the recording button on the interface, and the transcription tool displays the text of the sentences they read. This allows the teacher and other students to assess the pronunciation of the student presenting in real time. After each presentation, the teacher collects the transcription text for review. The teacher may present a sentence or word on the classroom screen for whole-class discussion, allowing students to complete their presentations while the transcription tool tracks pronunciation accuracy. After class, students can review and practice using the transcription text. This design aims to create an engaging and effective learning environment by combining cooperative learning with technological support to boost students' confidence in speaking.

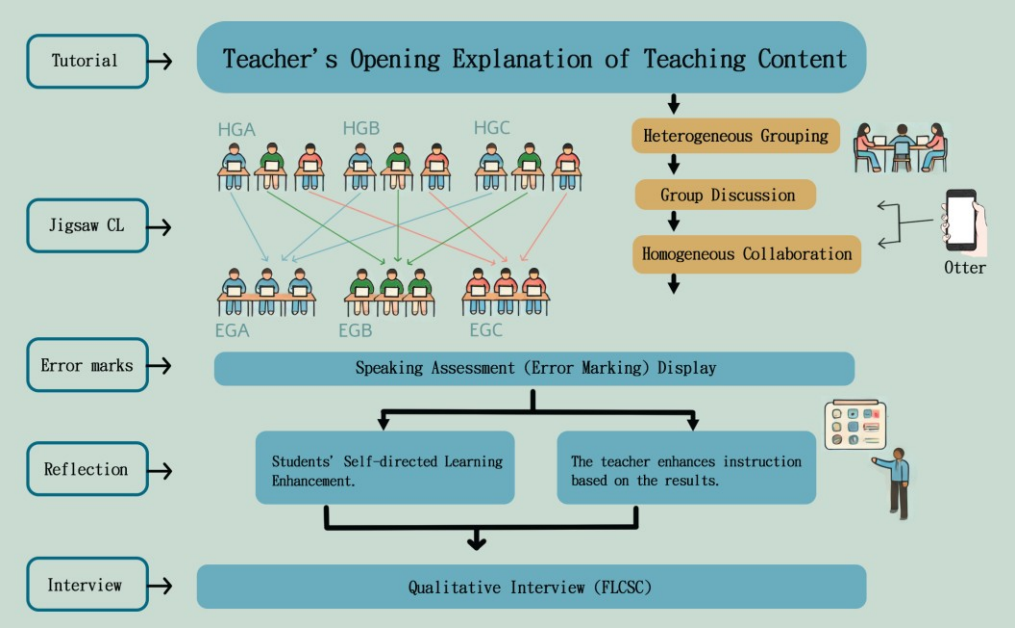

Figure 4: Flow of learning activity integrating with the meeting transcription tool.

4 Data Collection and Results Discussion

To verify the activity planning and relevance of this study, the following steps were taken:

1. Participants in this study included seven master's students and one fourth-year undergraduate student. Among them, five had prior experience in English-speaking courses, while three had limited exposure to spoken English practice. Their English proficiency levels varied, with CEFR levels ranging from B1 to C1, based on self-reported IELTS and TOEIC scores. The group consisted of five male and three female students, ensuring a balanced perspective on learning experiences.
2. Class Activities Implementation: Implement the study's class activities to ensure all eight students complete the corresponding tasks.
3. One-on-One Interviews: Conduct one-on-one interviews to deeply understand students' experiences and perceptions of using the meeting transcription tool for

improving their English-speaking. The entire interview process were recorded for documentation.

4. Feedback Survey: In addition to the FLCSC questionnaire, conduct an in-depth feedback survey to understand students' opinions on the operation of the meeting transcription tool and its assistance in English-speaking in class. Explore the tool's actual application, identify any issues students may encounter during use, and collect suggestions.
5. Subsequent Research Improvement Suggestions: Based on student feedback and opinions, adjust the research and activity planning to enhance the system's practicality and effectiveness.

This iterative process of testing and modification refines the research for improved usability and effectiveness. Due to the limited sample size of only eight respondents in my survey, I will refrain from discussing the statistical results, as they may not provide a reliable basis for conclusions. However, I conducted follow-up interviews based on the survey content, which adds significant research value. These interviews allowed for a deeper exploration of the participants' perspectives and experiences, offering qualitative insights that can enrich our understanding of the topic at hand. This approach emphasizes the importance of qualitative data in research, particularly when quantitative results are constrained by small sample sizes. It also provides valuable insights into students' experiences with the meeting transcription tool, guiding the direction of future research.

4.1 Past English-speaking Learning Experiences

Students generally report that the English education they receive in school provides limited opportunities for speaking practice, leading to a lack of confidence when using English in formal situations. To boost their speaking confidence, students frequently turn to cram courses, join English speech associations, or even live in English-speaking countries. This imbalance in English education can result in poor oral expression and a lack of speaking confidence, creating a vicious cycle.

To ensure anonymity, participants were assigned codes (e.g., M01, M09, M14) instead of using their real names. Each code corresponds to a specific student, allowing for structured analysis of individual feedback while maintaining participant confidentiality.

M13: "In elementary school, I spoke English more often in cram schools, but in junior and senior high school, the teaching became more exam-oriented, and there was less English-speaking."

M01: "In junior and senior high school, the focus was on academic content, and there was hardly any English communication."

M22: "There was less speaking in school. I only did presentations once or twice a semester. But I also joined an English speech association, where I shared once a week."

In English classes, teachers play a crucial role in providing feedback on students' speaking performance. However, traditional methods that focus on correcting grammar mistakes during every conversation can create a culture of constant evaluation. This emphasis on correction, exams, and grading may erode students' confidence and even lead to an aversion to the language. Language is a tool for communication, and the fear of speaking incorrectly can hinder this

purpose. Therefore, we should explore ways to adjust teaching methods, helping students use English with greater confidence.

M13: "After the speaking test, the teacher said she really liked my English accent, and I felt encouraged."

M14: "I wasn't particularly encouraged, and I even had some resistance to English."

4.2 Speaking Anxiety in English Presentations in Class

Research shows that second language speaking anxiety often stems from others' evaluations of the speaker, which is why speaking to teachers or superiors can be more nerve-wracking than speaking to peers. This highlights the importance of interaction among students, rather than one-way lecturing. Through interviews in this study, we gained valuable insights into the role of peer interaction. In a cooperative learning environment, students can observe each other's strengths and weaknesses, leading to mutual learning, healthy competition, and a shift in their mindset to speak English confidently. Future researchers should consider these findings and encourage students to ask questions and offer support in such environments, fostering positive changes in learning attitudes and promoting progress through mutual growth.

M14: "At the beginning, I was very nervous about speaking on stage. I would read many words quickly, and I felt the error rate was quite high. After several weeks of speaking and listening to other students' speeches, I began to slow down my pace and gradually grasp a reading rhythm that suited me."

M14: "The experience of speaking on stage has gradually enabled me to adapt to speaking and sharing in front of everyone. I clearly feel more courageous and confident in speaking, sharing, and asking questions than before."

4.3 Using the Meeting Transcription Tool Otter.ai for Learning English-speaking

Establishing correct self-perception is crucial in teaching. Language learning is a continuous process of correction and development, similar to acquiring a native language. Learners of Chinese or English may face situations where their pronunciation or word usage deviates from the standard. However, through sustained learning, imitation, and correction, they can develop accurate communication skills. This course encourages students to use technology-assisted tools like Otter.ai to practice speaking, receive objective feedback, and refine their pronunciation. This approach allows students to adopt a natural method for mastering language learning.

Otter.ai played a crucial role in the collaborative learning environment of this study. Through its automatic transcription feature, students could review their conversations after group discussions and analyze their pronunciation and expression. Additionally, Otter.ai’s built-in speaker identification function allowed students to compare their language performance across different discussion stages, enhancing their self-monitoring abilities. In this way, students not only improved their pronunciation accuracy but also gained greater confidence in speaking during the collaborative learning process.

M09: " Otter.ai provides a valuable reference for English speaking practice, even though it's not perfect. Being able to review the audio recordings allows me to analyze my pronunciation and pinpoint the root causes of my speaking challenges.

M13: "You can use Otter.ai to repeat pronunciations, confirm whether your pronunciation is correct, and recognize whether it is correct."

Otter.ai's purpose is to accurately record pronunciation, but its recognition results may not always meet users' expectations due to factors such as the recording environment, user accent, and operating method. Some students also question the need for Otter.ai, as they are more accustomed to using mainstream search engines.

M22: "The accent problem is therefore judged as another word, which also makes me think that if people with accents use speech recognition software, they may have the same problem."

M13: "Currently, I won't use Otter.ai to practice speaking because it takes less time to directly open Google. But will I have the opportunity to use it in the future? I think it's unlikely that I won't. It's just that I haven't encountered the need for it yet."

4.4 Reflection on English-speaking Pronunciation

This teaching approach promotes active, autonomous learning and metacognition. By using the Otter.ai meeting transcription tool, students can independently review and correct their pronunciation. Reading aloud with Otter.ai not only provides instant playback but also encourages deeper self-reflection as students compare their recognition results. Designed to meet students' learning needs and teaching objectives, this method aims to enhance their English expression skills while fostering autonomy and metacognition, aligning with contemporary educational principles.

M01: "In many places where sounds like s, t, d, and f require air sounds, or where the tongue needs to be spit out or the mouth needs to be opened wide."

M05: "Mainly, light sounds at the beginning of sentences are easy to be pronounced as heavy sounds, such as c/g, w/g, st/sl, cha/tra; and changes in word endings are not pronounced clearly, such as s, ed, sion, ing; there are also situations where the m sound is not closed tightly."

M09: "Some word connections between words are detected as a single word, for example, 'with the' is recognized as 'wasteful'."

M15: "Reading too quickly can lead to poor recognition of consonants such as t, s, z."

M22: "The pronunciation of word endings is not accurate, such as th, ness."

Summarizing the students' pronunciation reflections, it is clear that reinforcing the concepts of the three sound types and practicing them is crucial, especially for beginners, as it helps establish correct pronunciation habits early on. Students aware of their pronunciation issues can benefit from teacher assistance or using technology tools for objective guidance, ultimately improving pronunciation and boosting confidence in oral expression [1]. These recommendations can be directly applied in teaching to help students master English pronunciation and enhance their language skills.

4.5 Suggestions and Reflections on This Course Arrangement

After completing six English-speaking sessions, students provided feedback on the course content and activities, such as their presentations. This feedback offers valuable insights into students' opinions and satisfaction with the transcription system used in the jigsaw learning model, serving as a foundation for further research and improvements. Some students reported that the learning model has helped improve their English speaking and positively influenced their learning approaches outside the classroom. This outcome aligns with the research's expectations, allowing students to discover new learning strategies and continually explore the most effective learning model for themselves.

M01: "Using this method of comparing and marking mistakes after recording, I believe that in the long run, this way of practicing speaking will be very helpful. I found that I would unconsciously compare the differences between correctly and incorrectly pronounced words, reflecting on my pronunciation mistakes."

M14: "When browsing literature or reading news online, when encountering English, I will try to understand their pronunciation, not looking up first and then pronouncing, but pronouncing first and then looking up, feeling that my oral learning mode has changed."

According to the results of the FLCSC (Foreign Language Classroom Speaking Confidence) questionnaire, six students reported a reduction in anxiety when speaking in class after using Otter.ai and showed a greater willingness to actively participate in speaking activities. Students also mentioned in interviews that after practicing pronunciation with Otter.ai, their frequency of speaking in group discussions increased, indicating an improvement in their confidence.

Some students have suggested modifications to the course arrangement. For example, one student proposed revealing the course text on the day of the lesson to better assess students' speaking abilities and provide a more authentic English-speaking environment. This change could also offer a more accurate evaluation of whether students' speaking anxiety or confidence has significantly improved, as measured by the FLCSC questionnaire. Future researchers may consider this suggestion to enhance classroom activity design, ensuring that students remain challenged and stimulated without becoming too accustomed to a specific teaching method.

M15: "Perhaps redistributing the course text on the spot could better reflect individual English-speaking abilities."

5 Conclusion

In today's globalized world, English is becoming an essential communication tool. However, English education in Taiwan faces several challenges. The traditional teaching model focuses mainly on grammar and vocabulary, neglecting a comprehensive approach to speaking, listening, reading, and writing. Additionally, the exam-oriented nature of education leaves little room for real-world application, leading to a lack of interest and practicality in learning English speaking. Moreover, since English is not an official language, students have few opportunities to use it in daily life, which hinders their confidence in spoken English and pronunciation. With the rapid advancement of information technology, new opportunities arise to address these challenges. Meeting transcription tools like Otter.ai allow students to engage in autonomous learning, overcoming time and location constraints. Students can improve

their English speaking and pronunciation through practical application within their own context. Additionally, integrating technology-assisted tools and adopting more interactive teaching methods and cooperative learning environments can significantly enhance students' English abilities and overall learning experience. The commitment of dedicated teachers and students remains crucial for progress. Looking ahead, innovative teaching methods will continue to advance the quality of English education in Taiwan, making learning more enjoyable and better preparing students for the challenges of globalization.

This study not only validates the effectiveness of Otter.ai in improving pronunciation but also reveals that its integration with cooperative learning significantly reduces students' speaking anxiety and enhances their confidence in English communication. Through interview data and questionnaire results, we observed an increase in students' proactiveness in spoken expression and a greater willingness to participate in class discussions. This indicates that technology-assisted learning not only enhances language skills but also brings positive emotional impacts.

Acknowledgement

I would like to express my sincere gratitude to all those who supported and contributed to the development of this research. This paper is an extended version of the paper presented at the AAI 2024 conference, which has been published in its conference proceedings. The research reported in this paper has been supported by the National Science and Technology Council under research project numbers NSTC 113-2410-H-024-003-MY2 and the Ministry of Education under teaching practice research project number PSK1133865 in Taiwan.

References

[1] A. Ammar and Y. N. Thai, "The interrelation between the perception and production of English monophthongs by speakers of Iraqi Arabic," *Pertanika Journal of Social Sciences and Humanities*, vol. 24, pp. 1–10, 2016.

[2] M. T. Apple, *The Big Five personality traits and foreign language speaking confidence among Japanese EFL students*. Temple University, 2011.

[3] M. Asadian, "The Effect of Phonetic Transcription of TV News on Iranian Intermediate EFL Learners' Pronunciation Accuracy at Word Level," *International Journal of Research in English Education*, vol. 7, no. 1, pp. 61–71, 2022.

[4] P. E. Blosser, "Using Cooperative Learning in Science Education.," 1993.

[5] J. S. Brown, A. Collins, and P. Duguid, "Situated cognition and the culture of learning," *1989*, vol. 18, no. 1, Art. no. 1, 1989.

[6] R. Clément, "Ethnicity, contact and communicative competence in a second language," in *Language*, Elsevier, 1980, pp. 147–154.

[7] H. Evcim and Ö. F. İpek, "Effects of jigsaw II on academic achievement in English prep

classes," *Procedia-Social and Behavioral Sciences*, vol. 70, pp. 1651–1659, 2013.

[8] R. M. Gillies, "Cooperative learning: Developments in research.," *International journal of educational psychology*, vol. 3, no. 2, Art. no. 2, 2014.

[9] M. Hafizhah, L. D. Wahyuni, and Y. Lubis, "Learn English Pronountation And Phonetic Transcription With Elsa Speaking Application: A Student Perception," *Jurnal Riset Rumpun Ilmu Bahasa*, vol. 2, no. 2, pp. 65–73, 2023.

[10] B. Irawan and M. A. Tampubolon, "USING PHONETIC TRANSCRIPTION TO IMPROVE STUDENTS'PRONUNCIATION SKILLS," *Edulingua: Jurnal Linguistiks Terapan dan Pendidikan Bahasa Inggris*, vol. 7, no. 2, pp. 1–12, 2020.

[11] J. M. Keller, "Motivational design of instruction," *Instructional design theories and models: An overview of their current status*, vol. 1, no. 1983, Art. no. 1983, 1983.

[12] J. M. Keller, "Development and use of the ARCS model of instructional design," *Journal of instructional development*, vol. 10, no. 3, Art. no. 3, 1987.

[13] J. W. Peirce, "PsychoPy—psychophysics software in Python," *Journal of neuroscience methods*, vol. 162, no. 1–2, Art. no. 1–2, 2007.

[14] I. S. Samad and A. Aminullah, "Applying ELSA Speak software in the pronunciation class: Students' perception," *Edumaspul: Jurnal Pendidikan*, vol. 3, no. 1, Art. no. 1, 2019.

[15] A. Sholeh and U. Muhaji, "Pronunciation difficulties encountered by EFL students in Indonesia: sebuah studi kasus pada mahasiswa kelas integrated course semester 1 fkip Bahasa Inggris Universitas Kanjuruhan Malang," *Jurnal Inspirasi Pendidikan*, vol. 5, no. 2, Art. no. 2, 2015.

[16] R. E. Slavin, "Cooperative learning," *Review of educational research*, vol. 50, no. 2, Art. no. 2, 1980.