Using AI-Powered Speech Recognition Technology to Improve English Pronunciation and Speaking Skills

Nutprapha K. Dennis
Ubon Ratchathani Rajabhat University, Thailand

Abstract

This study aimed to investigate the impact of AI-powered Speech Recognition Technology (AI-SRT) in improving English pronunciation and speaking skills among EFL learners. Additionally, it explored the opinions and responses of EFL learners towards the use of this technology for pronunciation and speaking skill enhancement. The research employed a pre-test/post-test design and a survey questionnaire with a multiple-choice rating scale and open-ended questions. The study included a sample of EFL learners who received instruction and practice using an AI-SRT program over a designated period. Data were collected through pre-test and post-test assessments to measure changes in pronunciation and speaking skills. A survey questionnaire was administered to gather participants' opinions and responses regarding the program's effectiveness and their overall experience. The data were analyzed using descriptive statistics to examine the central tendency, variability and distribution of pre-test and post-test scores. Comparative analysis was conducted using paired t-tests to determine the significance of the improvement in pronunciation and speaking skills. Qualitative data from the survey questionnaire were analyzed thematically to identify recurring themes and patterns in participants' opinions. The findings of this study contribute to the understanding of the impact of AI-powered speech recognition technology on EFL learners' pronunciation and speaking skills. The results reveal the effectiveness of such technology and provide valuable insights into learners' perceptions and experiences. This research has implications for language educators, curriculum designers, and developers of educational technology in designing effective pronunciation and speaking skill enhancement tools for EFL learners.

Keywords: AI-powered speech recognition technology, EFL, English pronunciation, English speaking skills, pre-test/post-test design, survey questionnaire

Artificial intelligence (AI) is starting to revolutionize English pronunciation and speaking. AI-powered speech recognition technologies have shown potential in assisting students in improving their speaking skills. These programs, which are software applications or systems that employ AI algorithms and machine learning, convert spoken language into text. They are capable of analyzing and comprehending human speech, enabling devices or applications to respond to voice commands, transcribe audio recordings, or provide real-time, speech-to-text services.

While Speechling, Duolingo, and Google Assistant are prominent examples of AI-powered speech recognition technologies, it is important to explore additional alternatives to furnish learners with a diverse array of tools and resources (Garca-Sánchez, Soler-Urzúa, & Nussbaum, 2020). This study delves into various AI-powered speech recognition programs that are beneficial for improving English pronunciation and speaking abilities.

The use of AI technology has become more and more popular in language teaching. A significant number of studies have been done on how AI-powered speech recognition systems improve speaking and pronunciation. AI-based pronunciation feedback has been demonstrated by Garca-Sánchez and colleagues (2020) to enhance non-native English speakers' pronunciation. Similar increases in language learners' speaking fluency were observed by Zou and colleagues (2021) when they used AI-powered speech recognition.

The current study focuses on the implications of employing the Speechling program, an AI-powered speech recognition technology for English as a Foreign Language (EFL) learners, especially regarding pronunciation and speaking. Speechling is an online platform to help individuals improve their foreign language skills. The application uses artificial intelligence (AI) to deliver individualized pronunciation and fluency feedback. Learners can practice with actual human voices while recording sentences and receive expert guidance. It uses spaced repetition algorithms to improve language skills and is completely free.

The study aims to bridge the research gap by evaluating the effectiveness of AI-SRT in English pronunciation and speaking practice. Factors such as user-friendliness, program accuracy, feedback mechanisms, and additional functionalities will be explored. This inquiry will aid educators and students in selecting optimal tools for their requirements (Chen & Duan, 2020). Moreover, the study will explore the advantages and impact of speech recognition tools driven by AI. By broadening knowledge and comprehension of AI-SRT programs, students can make informed decisions in selecting the most suitable one for their personal learning objectives (Jia et al., 2019). Finally, this study seeks to empower language learners, educators, and researchers to harness AI technology effectively for practicing pronunciation and speaking skills. By evaluating AI-SRT programs for English improvement, learners can discover accessible language learning solutions tailored to their preferences. The ultimate goal is to assist students in achieving superior English speaking and pronunciation capabilities.

Conceptual Framework

The theoretical underpinnings of this research were grounded in well-established theories of language acquisition, second language learning, and computer-assisted language learning (CALL). To inform the design and implementation of AI-powered speech recognition technology (AI-SRT) programs for English pronunciation and speaking instruction, the study drew on Krashen's Input Hypothesis (Krashen, 1981) and Vygotsky's Sociocultural Theory (Vygotsky, 1934). These frameworks provided a robust theoretical base for understanding the cognitive and social dynamics of language learning and the advantages of integrating technology in language instruction. By incorporating these established theoretical perspectives, the study set a strong foundation for evaluating the effectiveness of the Speechling program and its impact on improving the pronunciation accuracy and speaking proficiency of language learners.

Background of the Study

Pronunciation and speaking present continual challenges for language learners, who have to navigate the complexities of linguistic and cultural integration to communicate accurately. The role of AI-powered speech recognition technology in language education has been a subject of interest, especially as it offers real-time feedback on pronunciation and speaking. Through the capabilities of machine learning and natural language processing, AI-SRT is able to pinpoint errors, mispronunciations, and areas in need of improvement within a learner's spoken language (Jia & Zhang, 2018). This advanced technology can facilitate personalized feedback, enabling tailored practice that enhances pronunciation, fluency, and overall speaking skills. Empirical evidence has suggested that AI-SRT could be instrumental in advancing language proficiency, with Computer-Assisted Pronunciation Training (CAPT) showing notable benefits for language learners (Jia & Zhang, 2018).

While existing studies yielded promising insights, there remains a need for further study into the application of AI-SRT in language learning. Additional study was essential to address potential limitations, refine its deployment, and verify its constructive contribution to language learning outcomes. This is important as its potency in augmenting pronunciation and speaking competencies for specific language functions has gained recognition in the realm of second language acquisition.

Moreover, the insights gleaned from such investigations have the potential to influence curriculum design and teaching methodologies. This study sought to delve into the influence of AI-SRT on the pronunciation and speaking abilities of language learners. Furthermore, it aimed to provide language educators and instructional designers with valuable guidance on how to effectively incorporate AI-powered tools in language education programs.

Statement of the Problem

The process of learning may have been impeded by limited access to high-quality learning tools and a shortage of experience between teachers with technology. These challenges may obstruct the effective practice of English pronunciation by EFL students. Therefore, the purpose of this study was to determine whether Speechling, an AI-powered speech recognition technology program, facilitated the accurate pronunciation of English by EFL learners. Furthermore, the study explored the manner in which EFL learners utilized an AI-SRT program to enhance their English pronunciation and speaking skills. This research supported educators in the implementation of a user-friendly and straightforward technology to enhance self-study, pronunciation, and speaking practice.

Research Questions

The study was designed to address two research questions:

1) What is the effectiveness of an AI-SRT program in enhancing EFL learners' English pronunciation and speaking abilities?
2) What are EFL learners' perceptions and responses to the use of AI-SRT for English pronunciation and speaking practice?

Purposes of the Study

This research is dedicated to exploring the potential benefits of utilizing an AI-powered speech recognition technology program to improve the English pronunciation and speaking proficiency of EFL students learning English. The study sets out to examine two principal areas.

The first area of focus is to assess the impact of the AI-SRT program on the participants' ability to enhance their English pronunciation and speaking skills. The research will employ systematic evaluation and in-depth analysis to determine the extent to which the AI-SRT program supports learners in the precise articulation of English words and phrases, as well as assesses its contribution to the overall advancement of their spoken language abilities.

The second aim is to explore EFL students' perceptions and reactions to the utilization of the AI-SRT program as a means to practice English pronunciation and speaking. By soliciting feedback from participants and conducting interviews and surveys, the study seeks to uncover learners' attitudes and opinions about the integration of AI-SRT technology into their language education. An understanding of their experiences will offer valuable insights into the program's receptiveness and the ease with which learners can engage with this educational technology.

Significance of the Study

This research expands the literature on using AI-powered speech recognition technology to improve English pronunciation and speaking for EFL learners. It aims to assist language instructors integrate such technologies into their teaching. The study will evaluate AI-SRT in language education to assist learners develop their language abilities without advanced learning resources. This study could further enhance English language instruction, improving communication between learners and native speakers. The study will examine AI-SRT's effectiveness for enhancing pronunciation and speaking skills, advancing English language learning. This could optimize language learning by using AI technology to improve teaching methods.

The findings could also assist curriculum developers develop more inclusive and adaptable learning materials for diverse learners. AI-SRT's practical benefits are demonstrated in the study to encourage policymakers to support the integration of advanced technologies in educational settings, which could lead to more language teaching innovation.

Literature Review

Importance of Pronunciation and Speaking Skills in Language Learning

Mastering pronunciation is essential for effective communication in a second language. Studies underscore the necessity of incorporating teaching of pronunciation, which includes elements such as individual sounds, intonation patterns, and speech flow, within language teaching programs to develop well-rounded oral abilities (Kang, 2016; Major, 2016).

Research conducted by Munro and Derwing (2016) and Saito and Lyster (2019) showed that phonetic training, activities that emphasize linguistic form, and constructive correction are key components in improving speech pronunciation. Furthermore, Thomson (2018) highlighted the need for dedicated instruction for educators in the art of teaching pronunciation, while also advocating for the integration of effective pronunciation instruction techniques into language education programs to aid in more successful language learning and oral proficiency.

Bohn and Munro (2020) delved into the difficulties that arise in learning how to pronounce words in a second language, shedding light on the obstacles students encounter and the targeted instruction needed to surmount them. Trofimovich and Isaacs (2019) argued for an increased emphasis on pronunciation within language teaching to enhance communicative skills. In addition, Zhang and Wei (2019) explored how second language acquisition, speaking abilities, and learner anxiety are interrelated, suggesting that fostering a relaxed learning environment can boost speech development. The collective findings of these studies make a compelling case that the role of pronunciation in language learning is pivotal, and requires personalized teaching strategies that build student confidence and skill. Recognizing the emotional aspects of language learning, such as anxiety, is also vital in crafting an educational setting that addresses personal learning hurdles and also advances overall language fluency.

Scholarly work emphasizes the importance of accurate pronunciation in second language acquisition and effective communication. To fully understand a language, teaching must account for sound subtleties, stress and intonation patterns, and speech flow. The relevance of these strategies across learning settings, and the ability to accommodate learners' diverse needs, such as native language influences, age, and unique cognitive approaches, vary and are subject to debate. Language teaching requires nurturing learning spaces and emotional support, as well as technical skills. Academics also want more pronunciation training for educators. Efforts to integrate pronunciation into academic curricula must be balanced against syllabus design, teaching materials, and educational institution strategy. Thus, while pronunciation is essential to language learning, its instruction requires overcoming logistical and learner-related obstacles.

Technology-Assisted Pronunciation and Speaking Training

Recent studies have shown a growing endorsement for the use of technology in the teaching of pronunciation and speaking skills for language learners. Research conducted by Morris (2020), Garca-Sánchez and colleagues (2020), and Liao and Xue (2019) have revealed the positive effects of artificial intelligence (AI) and Computer-Assisted Pronunciation Training (CAPT) in advancing learners' abilities to pronounce accurately and speak more proficiently. These technological tools, especially noted for their role in the betterment of vowel sounds and general spoken language performance, underscore the benefits of integrating digital aids into language education.

However, the success of technology-assisted pronunciation and speaking training is influenced by a variety of factors. Chen and Duan (2020) and Zou and colleagues (2021) highlighted the critical role of learner motivation, engagement, and individual learning differences in the efficacy of these programs. The literature suggests that learners' personal preferences and responses to training significantly shape the outcomes. Accordingly, CAPT programs that offer personalized feedback and adaptive learning experiences are recommended to cater to the diverse needs of learners. These adaptive features can optimize technology-assisted training by aligning with learners' attitudes, motivational levels, and specific requirements, thereby enhancing the overall effectiveness of such educational technology.

The literature reviewed herein supports using AI and CAPT to improve language learning pronunciation. However, a thorough analysis reveals that their efficacy depends on various factors. These factors include software quality, training program design, integration into comprehensive curricula, and adaptability to diverse learner needs and styles. Personalized systems require resources, and learners and instructors may have different technological skills. Although technology provides consistent feedback, it may lack the detail of human instructors. To improve language acquisition, technology-based language training must be carefully weighed against individual variations, learner attitudes, and instructor responsibilities.

AI-powered Speech Recognition Technology for Language Learning

Research examining AI-SRT suggests they offer substantial benefits for language learning. These technologies have been shown to enhance learners' pronunciation, fluency, and alleviate their anxiety towards language learning. For example, Li and Li (2021) affirmed that AI-driven language learning tools significantly improve pronunciation skills. Furthermore, Zhao et al., (2021) concluded in their meta-analysis that AI-assisted tools boost pronunciation precision, fluency, and reduce anxiety in individuals learning a second language. Yalcin and Korkmazgil (2021) discovered that an AI-infused mobile application markedly improved English pronunciation for EFL learners Kim (2019) observed that personalized AI-based pronunciation lessons notably advanced the English pronunciation of Korean EFL students. Additionally, Fan and colleagues (2019) reported improvements in the spoken English abilities of Chinese EFL learners through the use of AI. Collectively, these findings point to AI-powered speech recognition as a promising tool for enhancing the pronunciation of language learners.

Critical analysis of the literature reviewed on AI-powered speech recognition technology in language learning reveals several key areas for further study. There is a lack of discussion about AI's potential drawbacks, such as reduced human engagement and accessibility issues. The analysis suggests studying how the technology affects different learner demographics. It emphasizes the need for detailed pedagogical strategies to integrate technology into language learning. To determine the credibility and long-term effects of AI-facilitated learning, the referenced studies' research methodologies must be carefully examined. Ultimately, it emphasizes the need for resources and help in integrating these technologies into the curriculum and the challenges of teaching educators to use them.

Recent Developments in AI-SRT for EFL Learners (2023-2024)

Recent years have witnessed significant strides in the field of artificial intelligence (AI), particularly in its application to language learning and speech recognition technologies. Thes following literature reviews examined studies from 2023 to 2024 that exploring the use of AI-powered tools to improve English pronunciation and speaking skills.

With a cohort of English language learners, Mohammadkarimi (2024) tested Listnr and Murf AI pronunciation aids. A mixed-method study found that students who used these tools over traditional teaching methods improved their pronunciation. Positive attitudes toward AI-driven aids included increased engagement and confidence. The researchers also found difficulties in feedback interpretation and capturing subtle pronunciation nuances, suggesting AI technology needs improvement (Mohammadkarimi, 2024). Du and Daniel (2024) conducted a systematic review of AI chatbots' impact on English speaking skills. Their analysis of 24 studies showed that AI chatbots are still in their infancy, requiring further research. The results suggested AI chatbots can boost learning, reduce speaking anxiety, and improve pronunciation. Because AI chatbots can improve learning, English teachers, chatbot designers, and researchers should investigate them (Du & Daniel, 2024). Persulessy and colleagues (2024) examined civil engineering students' perceptions of learning engineering English vocabulary with ELSA

(English Language Speech Assistant). Students loved learning with AI, according to the descriptive research. The app's feedback mechanism motivated users, indicating a bright future for AI-assisted language learning in specialized fields (Persulessy et al., 2024).

The studies reviewed provide compelling insights into the integration of AI in language learning, particularly in enhancing English pronunciation and speaking skills. Mohammadkarimi (2024) presented a robust case for AI pronunciation tools, with quantitative data to back up qualitative perceptions. However, the study's limitations in feedback accuracy pointed to the need for more sophisticated AI algorithms capable of mimicking human-like feedback. Du and Daniel's (2024) systematic review broadened the scope of the field by analyzing the potential of AI chatbots across various studies. This meta-analytical approach was helpful in understanding broader trends and collective outcomes. Nonetheless, the review underscored the infancy of research in this domain, suggesting a vast potential for future studies to bridge the gap in knowledge and application. The focus on specialized vocabulary learning in Persulessy and colleagues' (2024) study added a unique dimension to the understanding of AI's role in language education. By examining a niche area of engineering English, this research offered a glimpse into the adaptability of AI tools across different disciplines. However, the study's sample size and single application focus might limit the generalizability of its findings.

Collectively, these studies highlight the transformative potential of AI in language learning. They also underscored the nascent state of this research field, advocating for continued development and exploration of AI capabilities. Future research should aim to address the challenges identified, such as feedback interpretation, and extend the application of AI tools to a wider range of learning environments and disciplines.

Research Methodology

To comprehensively evaluate the effectiveness of the Speechling program—an AI-powered speech recognition technology—in enhancing English pronunciation and speaking skills, this study employed a mixed-methods approach. Quantitative data was collected using a pre-test/post-test design to quantitatively measure the program's impact. Concurrently, qualitative insights were gleaned from participants' responses to a survey questionnaire, tailored to capture their perceptions of the program.

Participants and Pilot Testing

Participants for the research were drawn from a local university's sophomore EFL students during the Academic Year 2022. A convenient sample of 25 students, aged between 18 and 22, was selected from the English Listening & Speaking 3 Course. They were identified as having an intermediate level of English proficiency. Before commencing the full study, a pilot test was conducted with a smaller, separate group of students from a similar demographic, to refine the research instruments. The participants for the pilot test were chosen based on their availability and similarity to the target study group. Following the pilot, adjustments were made to the test

items and survey questions based on the initial feedback to enhance clarity, relevance, and engagement.

Research Instruments: Pre-test, Post-test, and Survey Questionnaire

The pre-test and post-test consisted of 30 items designed to evaluate pronunciation (20 items) and speaking skills (10 items). The Speechling program's content informed the test items, guaranteeing their pertinence. Each component of the test had a 30-minute completion time, with consistency in difficulty maintained across pre-test and post-test items.

The survey questionnaire, comprising 12 questions with a 5-point Likert scale and an open-ended question, solicited students' subjective experiences with the Speechling program. The selection of questions was informed by the pilot test feedback to ensure they effectively captured a comprehensive range of student experiences.

Data Collection Procedures

Data collection commenced with a thorough review of the research instruments. Participants information and privacy were safeguarded through strategic planning and the establishment of clear protocols. This included the distribution of consent forms, which were securely stored along with participants' data. Only the researcher had access to the collected data to maintain confidentiality.

During the first class, participants were introduced to the Speechling program and trained on its usage. The pre-test was administered under standardized conditions, followed by the intervention period with the Speechling program. Upon completion of the intervention, the post-test and survey questionnaire were administered with an emphasis on honest and thoughtful completion.

Data Transcription, Coding, and Analysis

Information gathered from the assessments and questionnaires was methodically transcribed and encoded. This entailed a thorough routine of inputting data, succeeded by statistical evaluation with the aid of suitable software tools. Descriptive statistics were utilized to encapsulate the data, and inferential statistics, including averages and standard deviations, evaluated the program's effect. The qualitative feedback underwent scrutiny to extract themes and discern patterns, and these insights were then examined in the context of the study's primary objectives.

Ethical Considerations

To protect participants, ethical protocols were strictly adhered to throughout the study. This included obtaining informed consent, ensuring anonymity in data handling, and securing data

access to authorized researcher only. These measures were paramount to ensuring the ethical integrity of the research.

In conclusion, the methodology was designed to be comprehensive and rigorous, incorporating pilot testing feedback for instrument refinement, adhering to ethical standards for participant protection, and employing a combination of quantitative and qualitative analysis to provide a holistic evaluation of the AI-SRT program's efficacy.

Results

Pre-Test and Post-Test

The step of the analysis was to delineate the outcomes of the pre-test and post-test evaluations. These evaluations were designed to measure learners' pronunciation proficiency. The data encompassed the performance of 25 participants

Table 1 illustrates the enhancement of pronunciation skills among 25 participants following an educational intervention. The mean score increased from 6.40 to 13.20. The pre-intervention score ranged from 10% to 70%, reflecting a range of abilities, while post-intervention scores ranged from 35% to 95%, marking substantial improvement. Despite a slight increase in the standard deviation from 3.14 to 3.30, the data suggests a more uniform improvement across the group, with median scores also rising from 35% to 75%. A paired t-test confirmed the statistical significance of these improvements, with a t-value of 7.46 and a p-value of less than .05. Hence, the progress in pronunciation abilities can be confidently attributed to the educational intervention.

Table 1
Pronunciation Abilities Pre-Test and Post-Test Results

Test	*n*	Mean	SD	Minimum	Maximum
Pre-Test	25	6.40	3.14	2.00	14.00
Post-Test	25	13.20	3.30	7.00	19.00

The data underscores an improvement in pronunciation among the 25 students from the pre-test to the post-test. This finding implied that educators and language instructors can enhance students' pronunciation skills by offering more speaking opportunities, employing visual aids, and addressing pronunciation errors promptly. A paired *t*-test was conducted to ascertain the significance of this improvement.

Table 2 displays a significant increase in the speaking skills of the 25 participants after an educational intervention. The means scores rose from 3.28 to 7.12 out of 20. The standard deviation remained relatively stable, shifting slightly from 1.55 on the pre-test to 1.47 on the post-test, indicating a consistent spread of scores around the mean. The minimum scores

increased from 3 to a higher minimum of 12, while the median scores also improved from 3 to 8, showcasing the participants' improved performance.

Table 2
Speaking Abilities Pre-Test and Post-Test Results

Test	*n*	Mean	SD	Minimum	Maximum
Pre-Test	25	3.28	1.55	3.00	7.00
Post-Test	25	7.12	1.47	8.00	12.00

A paired t-test confirmed the statistical significance of these improvements, with a t-value of 13.11 and $p<.05$. Hence, the progress in speaking abilities can be confidently attributed to the educational intervention.

General Implications and Limitations

The findings from this study offer valuable insights for language instructors aiming to bolster speaking proficiency. The successful intervention led to a demonstrable improvement in students' speaking skills, as shown by both descriptive and inferential statistical analyses.

However, it's important to recognize the limitations of this research. Conducted with a relatively small cohort, the study focused narrowly on a single aspect of language learning. Future research should consider larger sample sizes and encompass additional language skills to build upon these findings.

Despite these limitations, the investigation highlights the potential efficacy of targeted language-improvement interventions, with a particular emphasis on the role of descriptive and inferential statistics in gauging their success.

Table 3 shows 25 students' improved speaking abilities, with normalized scores rising from 0.33 to 0.70. The intervention improved participants' performance more uniformly, lowering SD from 0.18 pre-test to 0.16 post-test. Adding AI-powered speech recognition to the curriculum improved academic performance. This technology, along with personalized instruction and immediate feedback, improves pronunciation and oral communication. The study confirms previous research on technology-assisted language learning. However, the small sample size and participant homogeneity may limit the results' applicability.

Table 3
Speaking Abilities Pre-Test and Post-Test Results

Test	*n*	Mean	SD
Pre-Test	25	0.33	0.18
Post-Test	25	0.70	0.16

Finally, using AI in language learning improves pronunciation and speaking. Language teachers can provide more engaging, individualized learning experiences that improve language acquisition by incorporating such innovations.

Results from the Survey Questionnaire

The survey results revealed a unanimous agreement that the program exercises significantly enhanced the participants' pronunciation, speaking abilities, and vocabulary acquisition. Specifically, 22 (88%) students strongly believed in the program's effectiveness in improving pronunciation, while the majority confirmed its benefits for speaking skills. This aligns with the findings reported by Shao and colleagues (2018). Shao and colleagues (2018) and Lee and Jang (2018), reported similar findings regarding speech recognition systems. All participants acknowledged the program's role in learning new words, which echoes Kuo and colleagues' (2017) findings on vocabulary development through speech recognition technology.

Furthermore, the program was found to be engaging and attention-grabbing. Twenty-three (92%) of participants reported sustained focus during exercises. Motivation to practice English was reported by 15 (60%) participants, supporting Liu and Yang (2019)'s discovery of speech recognition's motivational impact. An overwhelming majority concurred that the program was instrumental in refining their pronunciation and bolstering their confidence in English speaking, a sentiment found by Wang and Hisao (2018).

Every participant strongly agreed on the program's efficacy in correcting pronunciation errors, reinforcing Zhang and colleagues' (2018) study outcomes. The program also proved helpful in amending speaking errors, as agreed by 18 (72%) participants, mirroring Lee and Jang (2018)'s findings. The exercises' variety and the flexibility of practicing English anywhere were highly praised, aligning with mobile and online language learning benefits highlighted by Kukulska-Hulme and Shield (2008). Additionally, the program's ease of use and enjoyment factor were emphasized, which are crucial for user satisfaction and effective learning as noted by Sanz-Torrent and colleagues (2018) and Jang and Wu (2018).

In summary, the Speechling program, facilitated by AI-powered speech recognition technology (AI-SRT), was unanimously recognized for its positive impact on language skills, with a strong endorsement from the participants for its pronunciation, speaking, and vocabulary exercises, motivational appeal, error correction capabilities, and enjoyable, accessible learning experience.

Quantitative Results

Participants were asked what they liked best about the Speechling program. The program received positive feedback from students for its effective approach to language learning. Key aspects that students valued the most included the program's flexibility, allowing for practice at any time and place, and its user-friendly interface, which enhanced the learning experience. The inclusion of a rich vocabulary and the ability to receive immediate, precise feedback on

pronunciation were highlighted as significant benefits, contributing to improved confidence and speaking abilities. The program's structure, which facilitates tracking progress, alongside a variety of accents for listening practice, was appreciated for helping students become accustomed to different English speakers and overcome language apprehension.

Moreover, the program's adaptive difficulty levels enabled personalized progression, and the engaging speaking practice questions kept learning enjoyable. Its high-precision speech recognition technology, which closely resembles human speech, added authenticity to the learning process. The program also offered a vast collection of audio recordings and realistic speaking scenarios, making it both comprehensive and practical. Affordability and easy access were mentioned as additional advantages for students who are mindful of their budgets.

In summary, the study highlighted the Speechling program's effectiveness in enhancing English pronunciation and speaking skills, supported by current research on the positive impacts of AI-assisted language programs. The program's varied features—flexible practice settings, diverse vocabulary, adaptive levels, and authentic speech recognition—contributed to a dynamic and effective learning journey. The study's qualitative findings suggests that technology-assisted pronunciation and speaking training should be an integral part of language-learning curricula, offering a tailored and enriched experience for learners. Further investigation is encouraged to explore the long-term benefits, best instructional practices, and integration strategies for such programs in diverse learning environments.

Discussion and Conclusion

The application of AI-powered speech recognition technology in enhancing English pronunciation and speaking skills among EFL students has promising outcomes. This study, underpinned by the interactionist approach to Second Language Acquisition, highlighted the significant role of meaningful interactions and targeted practice facilitated by AI-SRT. The instant feedback and the opportunity for repetitive, focused oral practice have proven to be effective in improving pronunciation and oral fluency, resonated with findings by Liu and Yang (2019) and Chen and Wang (2017). The technology's role as a virtual language instructor offering personalized feedback suggested its potential to substantially improve sound production and speaking skills.

EFL students' English pronunciation and speaking skills have improved with AI-SRT. This study, based on the interactionist approach to Second Language Acquisition emphasized AI-SRT-facilitated meaningful interactions and targeted practice. Liu and Yang (2019) and Chen and Wang (2017) found that instant feedback and repetitive, focused oral practice improved pronunciation and fluency. The technology's role as a virtual language instructor providing personalized feedback suggested it could greatly improve sound production and speaking.

Learner variables like motivation, attitudes toward technology, and prior language proficiency affected engagement and perceptions of the technology's impact. This is consistent with the work of Kim and Lee (2019) and Sung (2018), that reported learner motivation and positive

attitudes toward technology improve language learning outcomes. Wang and Sun (2020) and Gao and colleagues (2019) agreed that AI-SRT's authentic conversational practice and immediate feedback support interactionist SLA theory and the importance of meaningful practice. Wu and colleagues (2018) and Zhang and team (2022) recommended seamless and purposeful AI-SRT integration into language learning environments that aligns with instructional design and pedagogical goals. The study recommended including speech recognition technology in the curriculum to aid language development, provided that teachers and students receive proper training and support, as noted by Lee and Jang (2018) and Zhang and colleagues (2022).

The current study supports AI-SRT's efficacy, but practical considerations for its integration into language curricula must be critically examined. To keep up with learners' changing needs and language acquisition, technology must be assessed regularly. Technology's potential as a crutch rather than a tool for learning raises the question of how to balance it with other teaching methods to promote independent language development. The accessibility and inclusivity of AI-SRT must also be assessed. The technology allows anytime, anywhere practice, but socioeconomic factors and student digital literacy may limit access. Additionally, AI-SRT's ability to accommodate different learning preferences and disabilities has not been extensively explored. Furthermore, the study's findings suggest that while immediate feedback is beneficial, it is the quality and specificity of the feedback that may determine the extent of improvement in pronunciation and speaking skills. Therefore, ongoing refinement of speech recognition algorithms to provide nuanced feedback is crucial for the sustained effectiveness of AI-SRT.

Finally, AI-SRT may continue improving language learning, but more nuanced, longitudinal studies are needed to determine its long-term effects and optimal integration strategies for diverse learning contexts. This includes studying how it affects students of different proficiency levels, learning styles, and cultures. In conclusion, the Speechling program and other AI-SRT tools offer innovative language learning methods, but their practical application requires careful consideration of pedagogical, technological, and learner-related factors to ensure that they complement traditional language instruction and serve a wide range of learners.

Conclusion

An AI-powered Speech Recognition Technology was tested to improve EFL students' English pronunciation and speaking. The study yielded important findings.

Drawing on the interactionist approach to second language acquisition, AI-powered speech recognition technology aided meaningful interactions and targeted practice. Students could receive instant feedback and practice pronunciation and oral fluency with the technology.
Technology integration in language learning was also shown to be important. Students were supported and engaged by speech recognition technology powered by artificial intelligence. The technology provided personalized feedback and instruction as a virtual language instructor, improving students' sound production and speaking skills.

The study also considered how learner variables affect intervention efficacy. Learners' engagement with AI-powered speech recognition technology and perceptions of its impact on language learning were influenced by motivation, attitudes toward technology, and prior language proficiency. The results of this study contribute to the existing literature on AI-assisted language learning and shed light on the potential benefits of AI-powered speech recognition technology for enhancing pronunciation and speaking abilities. When designing effective language learning interventions, the study highlights the importance of considering the dynamic relationship between language acquisition theories, technology integration, and learner factors.

This study supports the use of speech recognition technology powered by artificial intelligence as a valuable tool for improving EFL students' English pronunciation and speaking abilities. The findings indicate that integrating technology into language learning environments can enhance the learning experience and result in more effective language acquisition. However, additional research is needed to investigate the long-term effects of these interventions and to address potential challenges associated with the use of technology in language education.

Limitation of the Study

This research study provided valuable insights, but its limitations should be considered in future studies. First, the study only included EFL students and may not apply to learners of other languages. AI-powered speech recognition may not work for all EFL learners due to age, language proficiency, and culture. Second, the small sample limited generalizability. A larger, more diverse sample could explain how AI-powered speech recognition affects pronunciation and speaking. Another drawback is self-reporting and subjective evaluations. Pre-tests, post-tests, and surveys are useful, but more objective measures are needed to confirm the findings. Expert evaluations and objective pronunciation tests would strengthen the study's findings.

The study also did not examine the long-term effects of AI-driven speech recognition on language learning. The technology's durability can be better understood through longitudinal student studies. The study did not address technological barriers to implementing AI-powered speech recognition in language learning environments. To ensure technology applicability and efficacy, future research should focus on system precision, user interface design, and accessibility.

This study illuminates how artificial intelligence-powered speech recognition technology may improve English pronunciation and speaking. To better understand how technology affects language learning, future research should address these limitations and expand on the findings.

Suggestions for Further Studies

This research study has a number of recommendations for future research that can be made in order to broaden and deepen our understanding of this field. These recommendations are based on the findings and limitations of the research study.

The study of AI-powered speech recognition technology in language learning includes multiple narrative threads with different perspectives on its effectiveness and use. Foundational longitudinal studies track student progress over time. These studies examine the long-term effects of AI-powered speech recognition technology on pronunciation, speaking, and language proficiency. By studying this perspective, researchers can better understand how this technology affects language learning over time and describe its benefits.

Comparative research seeks understanding. The user compares AI-powered speech recognition technology to other language learning methods to determine its benefits. Teachers learn how to use this technology most effectively by analyzing different teaching methods and intervention strategies. In complex studies, learner factor shifts are important. Researchers examine how age, language proficiency, motivation, and AI-powered speech recognition technology affect learner performance. This study illuminates the complex interaction between individual variations and technological intervention, resulting in customized and flexible learning environments for a variety of learners.

The study emphasizes technological optimization. AI-powered speech recognition technology is researched to improve its design and functionality. By assessing system clarity, user interface layout, and user input, they strategically plan to improve the technology's usability and effectiveness to meet learners' needs. Language acquisition is fascinating and varied when studied across cultures. Researchers test AI-driven speech recognition technology in different cultures and languages. This study examines how subtle cultural influences affect technology use and effectiveness by comparing the experiences and results of EFL students from diverse cultural backgrounds. This research helps us understand multicultural technology use.

These recommendations can help researchers better understand how AI-powered Speech Recognition Technology can improve EFL learners' English pronunciation and speaking skills.

References

Bohn, O.-S., & Munro, M. J. (2020). Second language speech learning: Theory, findings, and problems. In T. S. Cardoso & J. N. Y. Levis (Eds.), *The Routledge Handbook of Second Language Acquisition and Phonology* (pp. 69–85). Routledge.

Chen, C.-M., & Wang, H.-W. (2017). A review of using speech technology in English pronunciation training. *Journal of Educational Technology & Society*, *20(2)*, 228–238.

Chen, G., & Duan, J. (2020). Artificial intelligence in language learning: A review of current applications and future potential. *Computer Assisted Language Learning*, *33(3)*, 257–274.

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230

Fan, Y., Huang, X., & Tan, S. (2019). The effectiveness of artificial intelligence technology in improving Chinese EFL learners' oral English proficiency. *Computer Assisted Language Learning, 32*(5–6), 495–514.

Gao, Y., Yang, J., & Zhang, X. (2019). The role of learner engagement in computer-assisted pronunciation training: A study with EFL learners in China. *Computer Assisted Language Learning, 32*(5-6), 507–529.

García-Sánchez, S., Soler-Urzúa, M., & Nussbaum, M. (2020). Effectiveness of an AI-assisted program to enhance English pronunciation in EFL university students. *Computer Assisted Language Learning, 33*(5–6), 568–590.

Jang, H., & Wu, Y. (2018). The impact of mobile game-based English learning on EFL students' motivation and English proficiency. *Journal of Educational Technology & Society, 21*(1), 1–13.

Jia, J., Wang, H., Yang, D., & Zhang, X. (2019). How effective is a computer-assisted pronunciation training (CAPT) program? A case study in China. *Computer Assisted Language Learning, 32*(5–6), 560–578.

Jia, R., & Zhang, C. (2018). The effectiveness of computer-assisted pronunciation training for speakers of different languages. *Educational Technology & Society, 21*(3), 261–274.

Kang, O. (2016). The importance of pronunciation in English communication. *International Journal of Language and Linguistics, 3*(1), 45–50.

Kim, J. (2019). The effects of artificial intelligence-based personalized pronunciation instruction on the English pronunciation skills of Korean EFL learners. *English Language Teaching, 12*(6), 32–51.

Kim, S., & Lee, J. (2019). The role of learner motivation and attitude in an intelligent computer-assisted language learning environment. *Computer Assisted Language Learning, 32*(5–6), 457–476.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon Press.

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL, 20*(3), 271–289. https://doi.org/10.1017/S0958344008000335

Kuo, C. H., Yang, Y. T. C., Chen, I. J., & Wang, T. I. (2017). Effects of using speech recognition-assisted multimedia pronunciation learning system on English pronunciation and vocabulary learning. *Journal of Educational Technology & Society, 20*(1), 243–256.

Lee, C., & Jang, J. (2018). Effect of mobile pronunciation technology on English pronunciation practice. *Journal of Educational Technology & Society, 21*(2), 187–198.

Li, Y., & Li, C. (2021). Using artificial intelligence to support language learning: A systematic review. *Journal of Educational Computing Research, 59*(5), 1295–1326.

Liao, X., & Xue, J. (2019). Computer-assisted pronunciation training (CAPT) on improving vowel production of Chinese EFL learners. *Computer Assisted Language Learning, 32*(7), 607–629.

Liao, Y. K., & Xue, S. (2019). Investigating the effects of computer-assisted pronunciation training (CAPT) on English vowel production of Chinese EFL learners. *Computer Assisted Language Learning, 32*(5–6), 579–604.

Liu, W. C., & Yang, C. F. (2019). Mobile-assisted speaking practice for EFL college students. *Journal of Educational Technology & Society, 22*(3), 203–217.

Major, R. C. (2016). *Teaching pronunciation: The reference complement*. Routledge.

Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Learning, 27*(1), 98–115. Sanata Dharma University. https://doi.org/10.24071/llt.v27i1.8151

Morris, R. L. (2020). AI meets CAPT: Artificial intelligence-supported computer-assisted pronunciation training. *Computer Assisted Language Learning, 33*(3–4), 349–377.

Munro, M. J., & Derwing, T. M. (2016). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. *Canadian Modern Language Review, 72*(2), 166–189.

Persulessy, S. I., Nikijuluw, R. P. J. V., & Jakob, J. C. (2024). Utilizing artificial intelligence in language learning: What about engineering students' perception? *Journal of English Education Program, 5*(1), 48–58. https://doi.org/10.26418/jeep.v5i1.71159

Saito, K., & Lyster, R. (2019). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Applied Linguistics, 40*(2), 244–268.

Sanz-Torrent, M., Moreno-Sandoval, A., & Blanco-Suarez, C. (2018). Design and validation of an innovative mobile application to promote language learning. *Educational Technology Research and Development, 66*(6), 1395–1411.

Shao, L., Gao, Y., Wang, L., & Li, Y. (2018). Effects of an online speech recognition system on learners' English speaking performance. *Educational Technology Research and Development, 66*(2), 427–445.

Sung, Y. T. (2018). Using a mobile app-based video reflection activity to enhance student motivation and achievement in speaking proficiency. *Computer Assisted Language Learning, 31*(5-6), 490–510.

Thomson, R. I. (2018). *Teaching and learning English pronunciation: Research and practice*. Routledge.

Trofimovich, P., & Isaacs, W. (2019). Second language pronunciation. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge Handbook of Second Language Acquisition* (pp. 349–369). Cambridge University Press.

Vygotsky, L. S. (1934). *Thought and Language* (E. Hanfmann & G. Vakar, Trans.). MIT Press.

Wang, J., & Sun, Y. C. (2020). Effects of focused instruction and interactional feedback on the improvement of L2 Chinese learners' Mandarin tones. *Computer Assisted Language Learning, 33*(5-6), 466–490.

Wang, S., & Hisao, C. (2018). Effects of the use of an automatic speech recognition system on Taiwanese EFL learners' oral proficiency. *Computer Assisted Language Learning, 31*(1–2), 39–59.

Wu, W. H., Wu, Y. C. J., Chen, C. Y., Kao, H. Y., Lin, C. H., & Huang, S. H. (2018). Review of trends from mobile learning studies: A meta-analysis. *Computers & Education, 123*, 119–135.

Yalcin, S., & Korkmazgil, S. (2021). The effectiveness of an AI-powered mobile application on the English pronunciation skills of EFL learners. *International Journal of Emerging Technologies in Learning (iJET), 16*(12), 135–148.

Zhang, H., Song, Y., & Lu, X. (2018). Investigating the effect of using a mobile-assisted pronunciation application on learners' pronunciation accuracy and fluency. *Language Learning & Technology, 22*(3), 111–129.

Zhang, Y., & Wei, L. (2019). The relationship between speaking ability and speaking anxiety in second language acquisition. *The Asia-Pacific Education Researcher, 28*(3–4), 181–192.

Zhang, Z., Li, Y., Wang, Y., & Cheng, G. (2022). The effectiveness of artificial intelligence technology on foreign language learning: A meta-analysis. *Educational Technology Research and Development, 70*(2), 713–741.

Zhao, Y., Chen, Y., & Ma, F. (2021). Effects of AI-assisted technology on second language pronunciation accuracy, fluency, and anxiety: A meta-analysis. *Computer Assisted Language Learning, 34*(7), 756–779.

Zou, D., Deng, Z., Li, C., & Chen, Q. (2021). Improving oral English skills with artificial intelligence: The effectiveness of AI-powered speech recognition technology. *Educational Technology Research and Development, 69*(4), 1919–1944.

Corresponding author: Nutprapha K. Dennis
Email: nootprapa@gmail.com