AI-Driven Pronunciation Assessment: The Impact of SpeechAce on EFL Learners' Pronunciation Competency

Nguyen Ngoc Vu[1], Vo Thiet Thach[1], Tran Vu Diem Thuy[1*]

[1]Ho Chi Minh City University of Foreign Languages - Information Technology, Vietnam
[*]Corresponding author's email: thuytvd@huflit.edu.vn
[*] https://orcid.org/0009-0006-6408-0152
https://doi.org/10.54855/callej.252635

ABSTRACT

The study examines the impact of SpeechAce, an AI-driven pronunciation and fluency assessment tool, on enhancing students' pronunciation skills. Utilizing a mixed-method experimental design, the research involved an experimental group using SpeechAce and a control group relying on traditional methods. Quantitative data were collected through pretests, post-tests, and questionnaires. Statistical analysis was conducted to compare the effectiveness of SpeechAce with conventional teaching methods. In-depth interviews were conducted to collect qualitative data on participants' experiences and perceptions. The results revealed significantly higher post-test pronunciation scores in the experimental group compared to the control group. Additionally, the "students perceived SpeechAce as a user-friendly and effective tool, particularly praising its immediate feedback and interactive features. Furthermore, SpeechAce's ability to pinpoint specific pronunciation errors improved students' performance. These findings have important implications for language educators and developers of educational technology, suggesting that integrating AI tools like SpeechAce can significantly enhance language acquisition and support the integration of technology in English language teaching and learning.

Introduction

Pronunciation is an essential language component that learners need to acquire in order to communicate effectively. Unfortunately, teaching pronunciation in Vietnam has been overlooked since the ELT in the Vietnam education system favors receptive skills over productive ones. The focus at the middle and high school levels tends to emphasize mainly on grammar, reading, and listening, while productive skills, particularly speaking and pronunciation receive less attention (Vu, 2016b). The curriculum is heavily centered on exercises that develop reading comprehension and grammatical accuracy, with little attention given to speaking or pronunciation. These two language components are also excluded from the tests. Moreover, pronunciation is taught very briefly in the audiolingual method, which

emphasizes repetition and mimicry, offers only a superficial approach to pronunciation without fostering a deep understanding or interest in the nuances of spoken English. Such a reptetitive approach also demotivates students' learning process. As a result, after seven years of learning English, from grade 6 to grade 12, it is very hard for students to produce even the simplest everyday conversation. This lack of focus on pronunciation has profound significant implications: it not only limits students' ability to communicate effectively in English, but also diminishes their overall motivation and enthusiasm for learning the language.

In contrast, an increased emphasis on pronunciation could potentially transform students' attitudes towards learning English. By making pronunciation a central component of language instruction, students could experience greater enjoyment and motivation in their language learning journey. The rise of computer-assisted language learning (CALL) offers promising new avenues for pronunciation instruction, particularly through the development of automatic speech recognition (ASR) technologies. ASR has become an integral part of daily life, embedded in virtual assistants like Siri, Google Assistant, and Cortana, which can understand and respond to spoken commands. This technology has also found applications in education, particularly in the realm of pronunciation training (Kumar & Renuka, 2023; Li, Deng, Haeb-Umbach, & Gong, 2015).

One prominent example of ASR in pronunciation training is the software Praat, which analyses human speech by generating detailed phonographic representations. These graphs visually depict the pitch contours of speech, allowing learners to compare their pronunciation with that of native speakers. By providing this visual feedback, Praat helps learners identify discrepancies between their speech and the target pronunciation, facilitating more precise practice (Hirai & Kovalyova, 2023; Jiang, Jong, Lau, Chai, & Wu, 2023). Similarly, other learning English software which integrates ASR in their features such as Rosetta Stone, Duolingo are also popular amongst English learners. Recently, there has been the presence of SpeechAce, an ASR program that compares users' speech to the native speaker's ones. Such ASR software like Praat and SpeechAce have paved new paths in seeking a better approach to teaching and learning pronunciation. Thus, further studies need to be done on software similar to those to introduce new methods in teaching pronunciation to students and make their learning of pronunciation more motivating and meaningful.

In the context of global communication, English has emerged as a predominant language, making proficiency in English a valuable skill. Pronunciation is a critical component of language learning, as it directly impacts intelligibility and communication efficiency. However, non-English-major students often face challenges in acquiring accurate pronunciation due to limited exposure and practice opportunities. At a private university in Ho Chi Minh City, catering to a diverse student body, many of whom are non-English majors, needs effective tools and methodologies to enhance English language proficiency across its curriculum. Nevertheless, non-English major students do not receive extensive pronunciation instruction in the classroom. Pronunciation only takes a small part in each unit in the book English File (by Oxford University Press). Moreover, formal pronunciation teaching is time-consuming because instructors must address individual students' varied pronunciation errors. As a result, pronunciation is often taught reactively—only in response to noticeable mistakes—or

sometimes overlooked if the errors do not hinder intelligibility. This reactive approach has not been effective, as students continue to repeat previously corrected pronunciation mistakes.

Given these challenges, ASR tools like SpeechAce could provide a valuable solution for this university. By integrating such technology, students would have access to personalized and consistent pronunciation practice, reducing the dependency on in-class instruction while enhancing their overall language proficiency.

Literature review

Pronunciation in Language Learning

Pronunciation is a critical component of language learning, playing a vital role in a speaker's ability to communicate effectively. According to Fraser (2000), clear pronunciation is essential for ensuring that a speaker's intended message is understood by listeners without ambiguity or misinterpretation. Correspondingly, pronunciation competency, as defined in the realm of language learning, entails the ability to accurately produce the sounds of English, including its vowels, consonants, and stress patterns, in a way that is intelligible to listeners (Celce-Murcia, Brinton, & Goodwin, 2010). Derwing and Munro (2005) emphasize the importance of intelligibility in pronunciation, highlighting that the primary goal is effective communication rather than imitating native-speaker norms. Therefore, English pronunciation competency is a multifaceted skill set crucial for successful communication in English, focusing on intelligibility, comprehensibility and interpretability rather than adherence to native-like pronunciation.

In the context of learning English as a second language, pronunciation greatly influences intelligibility, which is the ease with which a listener can comprehend spoken words (Derwing & Rossiter, 2002). Even when a learner's grammar and vocabulary are accurate, poor pronunciation can lead to misunderstandings, thereby undermining communicative competence, which is the ultimate objective of language acquisition (J. M. Levis, 2005). Besides, acquiring accurate pronunciation poses several challenges for language learners, particularly those whose native language differs significantly from English. One of the primary obstacles is the phonetic disparity between the learner's first language and English. For instance, Vietnamese has a more limited range of vowel sounds compared to English, making it difficult for Vietnamese learners to distinguish and produce English vowels correctly (Vu, 2024). This phonetic gap can create persistent pronunciation errors that are hard to overcome without targeted instruction and practice.

Furthermore, many learners suffer from a lack of exposure to native English speakers and authentic spoken English. This limited exposure hampers their ability to internalize the subtle nuances of pronunciation and develop a natural-sounding accent (Gilakjani, 2012). The situation is exacerbated by the fact that pronunciation is often neglected in educational curricula, as is the case in Vietnam, where traditional language teaching methods prioritize grammar and reading over speaking and pronunciation (Vu, 2016b). When pronunciation is taught, it is frequently done using outdated methods that fail to engage students or address their individual pronunciation challenges effectively (Derwing & Rossiter, 2002). Psychological

factors also play a significant role in the acquisition of pronunciation. Learners often experience anxiety or embarrassment when speaking in a foreign language, particularly if they are unsure of their pronunciation. This anxiety can lead to a lack of confidence and reluctance to practice speaking, which further impedes their progress (Horwitz, 2001). Moreover, learners tend to transfer the pronunciation rules of their native language to English, resulting in errors. For example, Vietnamese speakers often struggle with the pronunciation of final consonant sounds in English, which are less prominent in Vietnamese (Vu, 2016a).

Pronunciation proficiency has a profound impact on overall language learning outcomes. Learners with good pronunciation are generally more confident in their speaking abilities, which encourages more frequent practice and participation in conversations (Derwing & Rossiter, 2002). This increased practice leads to greater fluency and competence in the language. Additionally, accurate pronunciation enhances listening comprehension, as learners become more attuned to the sounds and patterns of the language (Field, 2005). Conversely, poor pronunciation can lead to frustration and discouragement, especially if the learner's spoken language is frequently misunderstood. This frustration can decrease motivation and engagement in language learning activities, ultimately slowing progress (Trofimovich & Isaacs, 2012).

Teaching and Learning English Pronunciation in Vietnam Context

In Vietnam's public schools, English lessons mostly focus on grammar and skills like reading and listening, while pronunciation tends to be overlooked. The current 7-year English program doesn't have specific sections for pronunciation practice—it usually just leaves students to learn on their own using cassette tapes. But without proper guidance in the classroom, that kind of self-study doesn't really work. Because of packed schedules and pressure to cover the curriculum, teachers don't often have the chance to focus on pronunciation unless a student is clearly struggling. On top of that, with big class sizes, old materials, and not enough time, it's tough for teachers to teach English effectively. Most students also rely heavily on the teacher's explanations and aren't used to learning independently, which makes structured support for pronunciation even more important.

Having realized the problem of the old curriculum, the Ministry of Education and Training launched the National Foreign Languages 2020 Project, hoping to help students develop all four skills equally and be able to use English to perform real-life tasks. Unfortunately, the project was not a success. The Minister of Education and Training admitted that Project 2020 failed to achieve its target. As acknowledged by the Minister of Education and Training, Project 2020 did not achieve its targets (T. Nguyen, 2017). Vietnamese students' English pronunciation remains largely unintelligible to both native English speakers and even other Vietnamese English speakers (H. N. Nguyen, 2023). This failure underscores the urgent need for greater attention to pronunciation instruction and the adoption of new methods that can better motivate students and enhance their communication skills. The lack of emphasis on pronunciation is compounded by the fact that English teachers in Vietnam are generally not well-equipped to teach this skill. A recent study by Loc and Newton (2020) revealed that most teachers approach pronunciation instruction in unplanned and reactive ways, largely because their own training did not include formal pronunciation classes. This was because they were not offered formal pronunciation classes during their training. As a result, students' errors were corrected with

certain repetitive techniques such as recasts or prompts. However, these types of feedback do not have a long-term effect on students' pronunciation competence. If overused, they might demotivate the students' pronunciation learning. They may also prevent students from developing their autonomy. Thus, the students may never be able to learn pronunciation on their own. The study also highlighted that teachers expressed a strong need for more formal pronunciation training sessions to develop the skills necessary for effectively teaching pronunciation to their students.

It is a common aspiration for learners of any skill to master all its components, and English learners are no exception. As Ha and Bao (2023) noted in a study, students often aspire to achieve native-like pronunciation or, at the very least, to be intelligible when speaking English. The study also identified the common difficulties students face in learning pronunciation and their desire for proper instruction that would help them express themselves more clearly and become more autonomous in their learning. These findings align with those reported by Nguyen (2017), emphasizing the importance that both teachers and students place on effective pronunciation teaching and learning. Ultimately, intelligible pronunciation is key to motivating students and enabling them to apply what they have learned effectively in communication. However, integrating pronunciation instruction into an already-packed curriculum presents a significant challenge. If there is insufficient time to include formal pronunciation lessons within the existing schedule, alternative approaches must be explored to teach pronunciation effectively without adding to the students' workload. This could involve the use of innovative tools and methods that can be seamlessly integrated into the current curriculum, thereby enhancing pronunciation instruction without overburdening teachers or students.

Automatic Speech Recognition for Pronunciation Teaching and Learning

As the name suggests, with automatic speech recognition (ASR), our speech is listened to and analyzed by a machine for various purposes. For example, we can give commands to our phones by talking to them using Siri for iPhones, Google Assistant for Android phones, or Cortana for Windows laptops. The machines listen to our speech, analyze it, and execute the command accordingly. ASR technology has also been applied in the field of teaching English pronunciation through Computer-Assisted Pronunciation Training (CAPT) programs. Learners record their utterances of words, phrases, or sentences into the CAPT program, which then provides feedback focusing on the phonological features of their utterance. The types of feedback and functions may differ depending on the programs. For instance, Spring and Tabuchi (2022) discuss how ASR-based pronunciation practice can aid learners in improving second language pronunciation, emphasizing the impact of treatment length and guided practice on specific pronunciation points. Similarly, Moxon (2023) highlights the advantages of ASR technologies, such as instant feedback and visual representation of sounds, which contribute to improved intonation and pronunciation development.

One of the first and most notable program that used ASR in training pronunciation is Praat. Developed by Paul Boersma and David Weenink of the University of Amsterdam, Praat analyzes a learner's speech in terms of intensity, pitch height, duration, or formants, then compares the speech with a native speaker's model through a display of speech waveforms. Research on the application of Praat in pronunciation instruction has yielded promising results.

For example, Gorjian, Hayati, & Pourkhoni (2013) found that using Praat helped learners acquire important prosodic features, including stress and intonation, which are critical for effective pronunciation. More recently, a 2019 study by Spauling and Breazeal explored the use of a pronunciation-based child robot, demonstrating that children's literacy and pronunciation skills improved after engaging in oral interactions with the robot. Similarly, ElsaSpeak, an English-speaking practice app that utilizes ASR, has received positive feedback from learners, who reported significant improvements in their pronunciation skills after using the app (Kholis, 2021; Sholekhah & Fakhrurriana, 2023). Despite the successes of ASR-based CAPT programs, they are not without limitations. A key feature of these programs is their ability to provide instantaneous feedback, which, while beneficial, also presents inherent drawbacks. There are primarily two types of feedback in ASR pronunciation programs: visual feedback and ASR feedback. Visual feedback, as provided by programs like Praat and Pronunciation Power, displays the learner's speech in a waveform graph, which requires specialized knowledge to interpret. Without guidance from a teacher, learners may struggle to understand and apply this feedback independently, limiting its effectiveness. On the other hand, ASR feedback is more straightforward and basic. In programs like Rosetta Stone and Duolingo, learners' utterances are simply marked as correct or incorrect, with the degree of intelligibility often represented by a visual indicator, such as a circle. However, these programs typically do not provide detailed analyses of the phonological features of the learners' speech, which could be crucial for addressing specific pronunciation errors (Bajorek, 2017).

While CAPT programs that utilize ASR have shown promising results, significant challenges remain. As J. Levis (2007) points out, "CAPT systems often suffer from difficulties in giving learners adequate, accurate feedback and an inability to provide an accurate and automatic diagnosis of pronunciation errors" (p. 185). In the following section, the discussion will focus on an ASR-based CAPT program that has been used to teach pronunciation, with the hope of addressing some of the gaps and limitations identified in previous programs.

SpeechAce: An Emerging Tool in Pronunciation Instruction

Given these challenges in existing ASR tools, SpeechAce has been developed to offer more targeted feedback and greater accessibility for learners.

SpeechAce, developed by ex-Microsoft veterans Ian Cheung, Abhishek Gupta, and Ahmed El-Shimi, is a speech recognition API designed for fluency and pronunciation assessment. Its primary feature is the ability to score a learner's speech in real time, identifying mistakes at the syllable and phoneme levels. The SpeechAce API supports a range of speaking activities, including word and sentence-level pronunciation assessments, passage-level fluency assessments, and the estimation of IELTS and PTE speaking scores (Moxon, 2021). Additionally, it can transcribe and evaluate spontaneous responses, as well as assess grammar and vocabulary. The SpeechAce API returns an overall score in each evaluated utterance and detailed sub-scores providing feedback on mistakes at sentence, word, syllable, and phoneme levels.

SpeechAce's functions are straightforward and user-friendly for both the teachers who design the tasks, and the users doing the tasks. First, teachers key in the prompt. After that, users turn

on the microphone on their electronic devices and record their answers. SpeechAce then analyzes the speeches and gives feedback on users' pronunciation. SpeechAce can be integrated into various Learning Management Systems (LMS), such as Moodle, Blackboard, Canvas, thus teachers can customize the tasks depending on the speaking skills they want learners to practice. Teachers can also listen to learners' speech and add their comments besides the comments from SpeechAce. Users can practice their pronunciation endlessly, and they can also have access to SpeechAce from different types of electronic devices. Additionally, the feedback is comprehensive and consists of two main parts. The first part is the percentage score from 0 per cent to 100 per cent on how close the user's utterance is to the model one of the native speakers. The percentage is also accompanied by comments such as "You got it! Are you a native speaker?", "You are not bad", "Your pronunciation is off". The second part is a detailed analysis of individual syllables and phonemes of users' utterances. Users can then compare their pronunciation with the models. In addition to general feedback, SpeechAce also provides IELTS scores for speaking, evaluating both fluency and pronunciation. Users can record their responses to a prompt, receive an IELTS score, and review detailed feedback on their specific mistakes, along with advice on improving their performance.

Figure 1

SpeechAce task and feedback
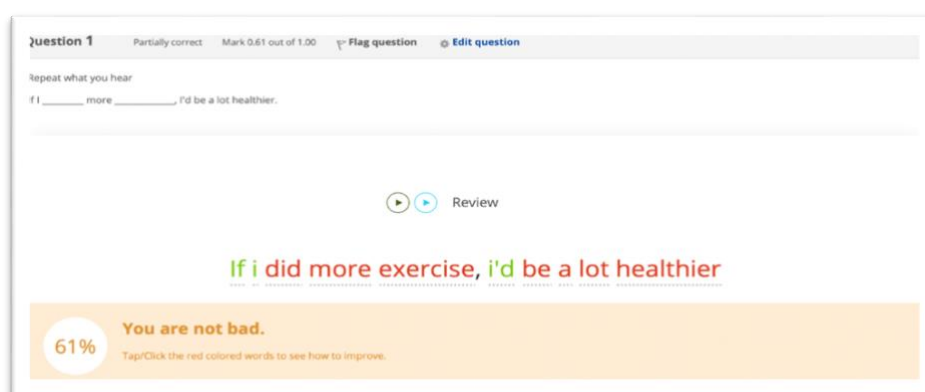
Compared to earlier ASR-based programs like Praat, Rosetta Stone, and Duolingo, SpeechAce offers more specific and actionable feedback. While programs like Praat and Pronunciation Power rely on complex waveform graphs that require specialized knowledge to interpret, and others like Rosetta Stone offer only basic correct/incorrect feedback, SpeechAce provides detailed phoneme-level feedback that pinpoints precise errors in pronunciation (Hassim, Kassim, Kassim, & Kassim, 2023). This level of specificity is essential for effective language learning, aligning with the pedagogical criteria suggested by experts like Pennington and Levis, who emphasize the importance of setting measurable goals, providing consistent and immediate feedback, and linking pronunciation to other aspects of communication.

Figure 2

Detailed analysis of pronunciation in SpeechAce
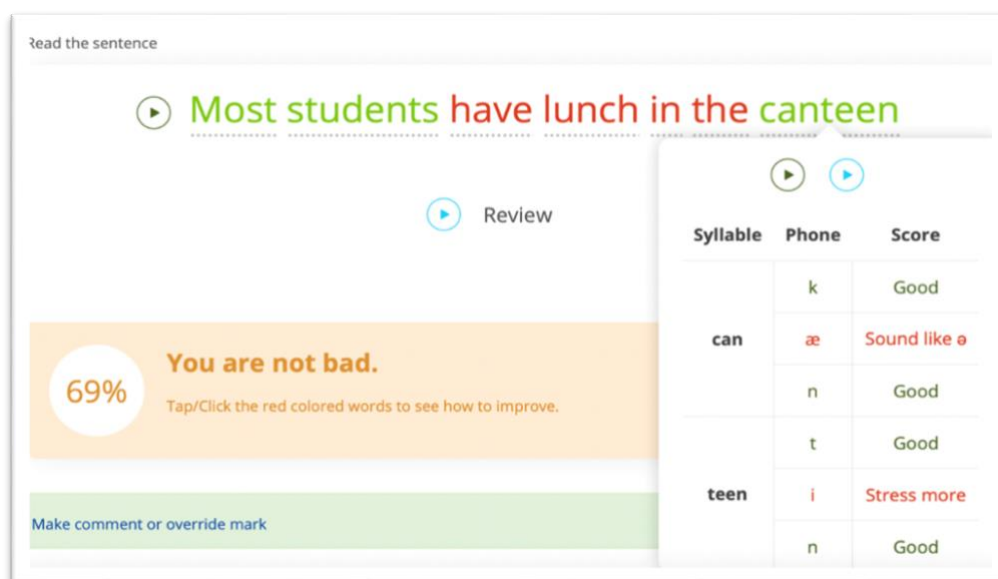
SpeechAce meets these criteria by offering pronunciation models in both American and British accents, providing percentage-based scores that help learners set clear goals, and integrating seamlessly with various LMS platforms. This flexibility allows teachers to design customized tasks, and detailed feedback enables students to practice and improve their pronunciation at their own pace, regardless of time or location. Although studies on the effectiveness of SpeechAce are still limited, the existing research suggests significant potential for the tool to enhance pronunciation skills. For instance, Moxon (2021) reported positive outcomes in learners' pronunciation accuracy and self-confidence, though both studies noted limitations such as the lack of control groups or balanced gender representation. Despite these limitations, SpeechAce has demonstrated its value as a resource for addressing English pronunciation challenges, particularly in educational settings.

Given the potential of SpeechAce and the gaps in existing research, this study aims to build on previous findings by including both a control group and an experimental group, with an equal number of male and female participants. By doing so, the study hopes to provide more robust evidence of effectiveness in improving pronunciation and to encourage further research into the application of ASR technology in pronunciation instruction. SpeechAce offers a promising solution for pronunciation training, with the potential to make pronunciation teaching more effective and accessible for both teachers and learners.

Research Gap

Despite growing interest in using Automatic Speech Recognition (ASR) tools like SpeechAce for pronunciation training, existing research remains limited in several key areas. Many prior studies have either lacked a control group or failed to ensure balanced participant demographics, making it difficult to draw definitive conclusions about the effectiveness of such tools. Additionally, while some studies have explored learners' improvement in pronunciation, few

have systematically investigated learners' perceptions of using ASR-based platforms in instructional settings. This study seeks to address these gaps by incorporating both a control and an experimental group, with an equal distribution of male and female participants, to offer more reliable data. Furthermore, by examining not only pronunciation outcomes but also learner attitudes, this research aims to provide a more comprehensive understanding of the impact and user experience of SpeechAce in pronunciation instruction.

Research Questions

To fulfill the purpose of the study, the survey aimed to answer the following research questions:

1. To what extent does SpeechAce improve students' pronunciation competency?

2. What are the students' perceptions toward learning pronunciation with SpeechAce?

These questions seek to explore the effectiveness of SpeechAce in enhancing pronunciation skills and to understand students' experiences and attitudes toward using this technology in their language learning process.

Methods

Pedagogic Setting & Participants

The research site is a private university in Ho Chi Minh City, established in 1991. It offers a diverse range of academic programs and serves a student body with varying levels of English proficiency.. The study was done at the Department of General English, which belongs to the Faculty of Foreign Languages. The department's main program is called English for International Communication (EIC). The program is designed to help students use English confidently and effectively in daily contexts, in multi-cultural working environment, and in academic settings. There are six levels in total, from EIC 1 to EIC 6. At the end of EIC 6, students are required to take the Aptis test, which is developed by British Council. Students have to obtain at least 120 points to successfully pass the level. The main course books throughout the program are the English File 3rd Edition (Oxford University Press), from elementary to intermediate plus. Students also have different supplementary materials depending on their levels to develop writing skills.

The participants of the study are from two EIC 4 classes coded EIC4 – 0500 and EIC4 – 0200. The two classes are chosen as intact groups for the study, in which, class EIC4 – 0500 (n=37) is the experimental group and class EIC4 – 0200 (n=37) is the control group. The demographic breakdown showed 33.8% male and 66.2% female participants. Both groups used the "English File – Intermediate" textbook (units 6-10) and attended three 2.5-hour sessions weekly. The experimental group practiced pronunciation using the SpeechAce platform, while the control group followed the textbook-based approach.

Design of the Study

To address the research questions, the study utilized a mixed-method quasi-experimental design. This approach, combining quantitative methods (pretests and post-tests) with qualitative methods (questionnaires and interviews), allows for a comprehensive evaluation of

SpeechAce's impact on pronunciation competency. The experimental design, involving control and experimental groups, establishes a causal link between SpeechAce use and pronunciation improvement. Although convenience sampling was used due to practical constraints, the approach remains effective in educational research. The combination of quantitative and qualitative data provides a holistic understanding of both the efficacy of SpeechAce and the students' perceptions of its use.

Data collection & analysis

The instruments used to collect data and assess the effectiveness of SpeechAce in enhancing pronunciation competency among non-English majors are elaborated, as follows:

SpeechAce-Based Pronunciation Task

Participants from the experimental group will receive pronunciation training with SpeechAce. To replicate the exact paper-based pronunciation task into SpeechAce is not possible. As a result, the researcher made use of all available features in SpeechAce to maintain the same characteristic between the SpeechAce-based pronunciation task and paper-based pronunciation task. Fortunately, SpeechAce is designed to integrate with the LMS system Moodle, hence there are more available features that can be adapted into SpeechAce to maintain its similarity with the paper-based pronunciation task.

Pre-Test and Post-Test

Both groups completed a written pretest and post-test to assess their pronunciation competency. The pretest comprised 37 questions from *Test Your Pronunciation* by Michael Vaughar (2002), focusing on phonemes, stress, intonation, and phonemic transcription. The post-test, consisting of 74 questions, was based on *English Pronunciation in Use* by Mark Hancock (2003). These phonological elements are fundamental to pronunciation assessment, as they form the basis of effective communication. The researcher graded both tests using the provided answer keys from the textbooks, making an inter-rater unnecessary.

Questionnaires and Interviews

The experimental group completed a 23-item questionnaire on a five-point Likert scale to gauge their attitudes towards pronunciation learning and using SpeechAce.

The study utilized standardized open-ended interviews with ten volunteers (2 males, 8 females) conducted individually in a classroom setting and recorded with their consent. Each four- to five-minute interview followed a consistent nine-question format. Participants, coded as S1 to S10, answered questions categorized into three sections: their background and challenges in learning pronunciation, perceptions of AI-based pronunciation training, and experiences with SpeechAce.

Results/Findings

The Impact of SpeechAce on Students' Pronunciation Competency

To assess the impact of SpeechAce on students' pronunciation competency, the study analyzed data from the pretest and post-test scores of both the experimental and control groups.

Pretest

Table 1

Descriptive statistics of pretest scores

Group	N	Mean	Standard Deviation
Control Group	30	30.93	5.95
Experimental Group	36	31.47	4.30

Table 2
Independent Sample T-Test of pretests results

Pretest	p-value	Sig. (2-tailed)
Equal variances assumed	0.167	0.671

An Independent Sample T-Test was conducted to ensure that both groups started with comparable pronunciation skills. As in table 1, the pretest mean scores were 30.93 for the Control Group (CG) and 31.47 for the Experimental Group (EG). The slight difference in mean scores (M = 30.93, SD = 5.95 for CG; M = 31.47, SD = 4.299 for EG) was evaluated for statistical significance. The results of the Levene's test (Sig. = .167) in table 2 confirmed equal variances, and the T-test showed a Sig. (2-tailed) value of .671, indicating no significant difference between the groups. Thus, both groups were deemed equivalent in pronunciation competency before the treatment.

Post-test

Table 3
Descriptive Statistics of Posttest Scores

Group	N	Mean	Standard Deviation
Control Group	30	51.8	5.839
Experimental Group	36	58.03	4.687

Table 4
Independent Sample T-Test of Posttests Results

Posttest	p-value	Sig. (2-tailed)
Equal variances assumed	0.257	0.00

As shown in table 3 and 4, the same analysis was performed on the post-test scores. The Control Group had a mean score of 51.80 (SD = 5.839), while the Experimental Group scored significantly higher, with a mean of 58.03 (SD = 4.687). The Levene's test again confirmed equal variances (Sig. = .257). The T-test results revealed a Sig. (2-tailed) value of .00, indicating a statistically significant difference between the post-test scores of the two groups. This

demonstrates that the Experimental Group, which used SpeechAce, showed a significantly greater improvement in pronunciation competency than the Control Group.

Students' Perceptions of Learning Pronunciation with SpeechAce

Results from questionnaires

The questionnaire was delivered to the students of the experimental group at the final session. It is used to gauge students' attitude and difficulties in pronunciation as well as their views towards adapting SpeechAce into their pronunciation learning. Initially, there were 24 items in the questionnaire, however, after the pilot test, one item was removed to satisfy the reliability and validity of the questionnaire. The questionnaire then consists of 23 items, divided into three sections.

The first section consists of 6 items to measure student's attitudes toward pronunciation and 1 item to survey students' preference in learning pronunciation. In the second section, students express their difficulties in learning pronunciation through a set of 7 items. The final section includes 8 items in multiple choice Likert scale form and one open-ended question. It surveys students' attitude towards using SpeechAce to learn English pronunciation. The questionnaire was developed on a five-point Likert scale, spanning from (1) strongly disagree to (5) strongly agree.

To ensure the questionnaire's reliability, Cronbach's Alpha was calculated for each section using SPSS version 28. The results showed acceptable reliability levels with Cronbach's Alpha values of 0.708, 0.723, and 0.755 for the three sections, respectively. These values indicate that the questionnaire was both valid and dependable, providing a solid foundation for analysing students' responses.

*Students' Perceptions and Difficulties in Learning Pronunciation*

Figure 3

Students' perceptions towards pronunciation
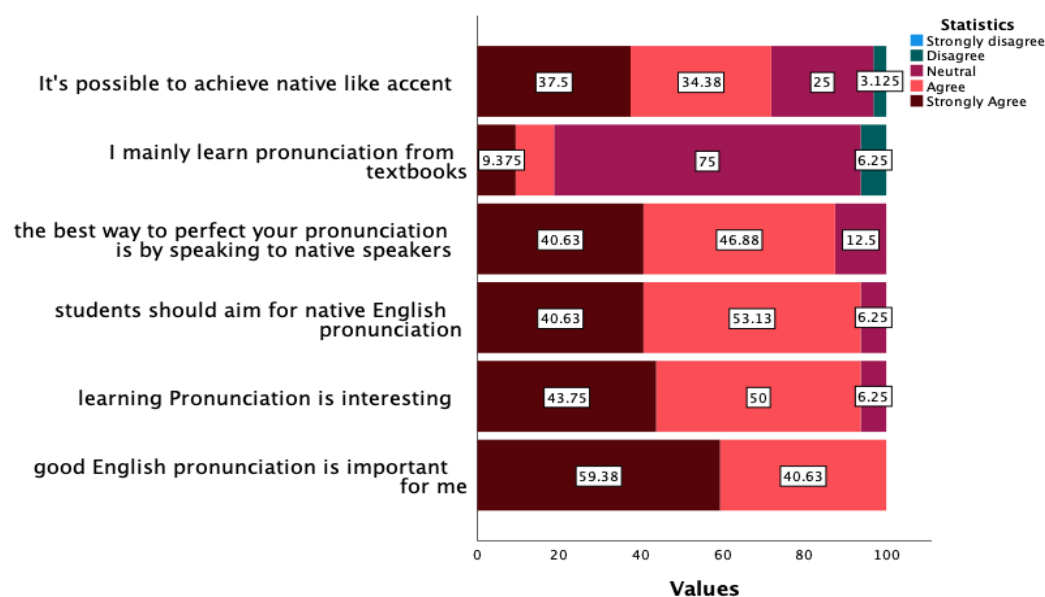
The first section (figure 3) revealed that all students agree that having good English pronunciation is essential for them, with 19 students (59.38%) strongly agree with this statement and 13 students (40.63%) agree with this. Most students showed their desire in learning pronunciation with half of the students surveyed found it interesting to learn pronunciation, while only 2 students chose neutral. From the first two items' results, students have very serious attitude towards learning English pronunciation, not only did they find it is an important component to acquire, but they also found interest in learning it. Most students target native-like pronunciation with more than half of the participants agree with this item and nearly half of them strongly agree with this, while only 2 students express a neutral stance and none of the surveyed participants strongly disagreed or disagreed with this item. The result in item 3 seems to synthesize with that of item 4, which ask participants to confirm whether they think the optimal way to perfect English pronunciation is by speaking to native speakers. The result shares similar pattern with item 3's results, with nearly 90% of the participants agreed or strongly disagreed, whereas no participants disagreed or strongly disagreed with this assertion. In item 6, a considerable number of participants viewed it possible to achieve native-like pronunciation. Only 1 student disagreed with this statement, meanwhile 8 of them took a neutral stance and the rest of the group either agreed or strongly agreed. It seems that students believe it is possible to speak English like a native speaker with proper practice and learning techniques. However, item 5's result has shown that students do not receive on-going pronunciation learning, with 24 participants chose neutral when being asked to confirm if textbooks are their main sources of learning pronunciation.

Figure 4

Students' difficulties in learning pronunciation
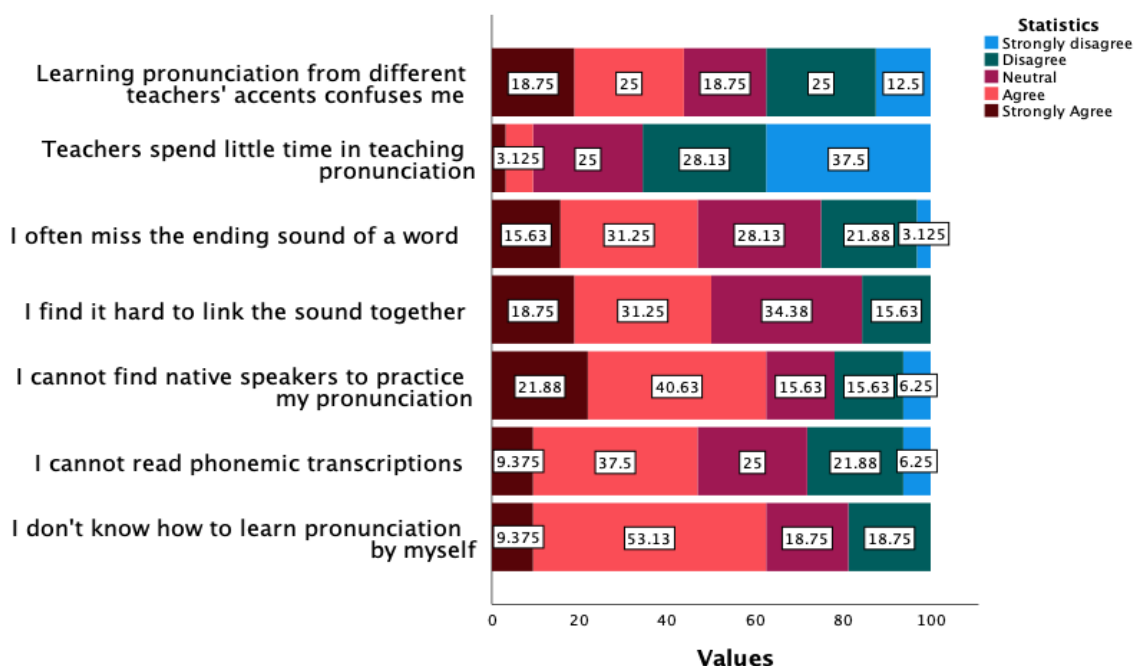
The next section explored the challenges students faced. As shown in figure 4, over half of the participants admitted they did not know how to learn pronunciation independently, with 17

agreeing and 3 strongly agreeing with this statement. Nearly half struggled with reading phonemic transcriptions, and many found it difficult to find native speakers to practice with. Ending and linking sounds were particularly troublesome, with a significant number of students agreeing that these aspects were challenging. Despite these difficulties, most students disagreed with the statement that teachers spend little time on pronunciation, and there were mixed views on whether learning from different teachers' accents caused confusion.

Students' Perceptions of SpeechAce

Figure 5
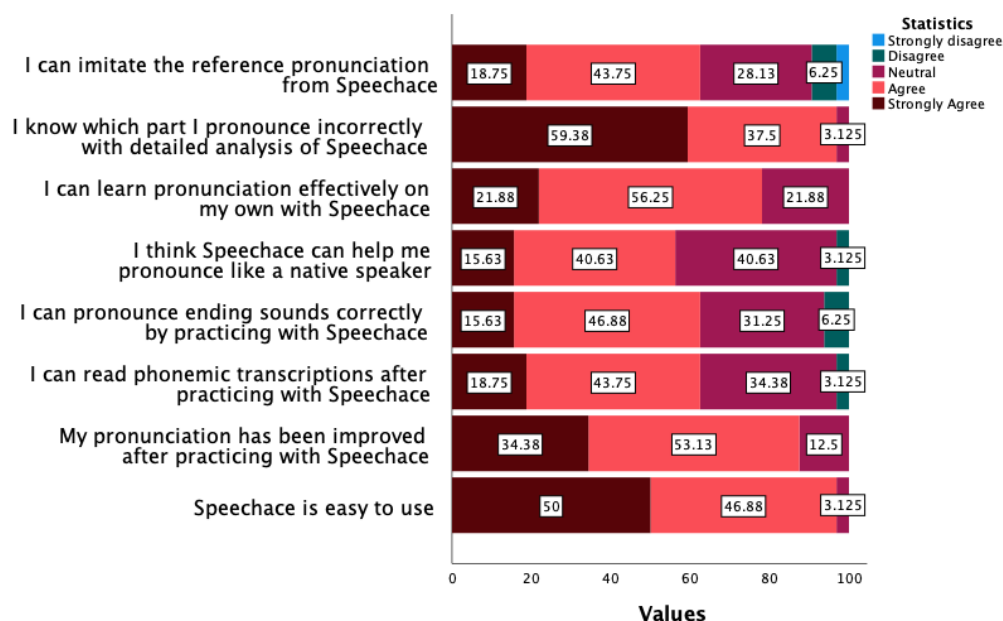
Students' perceptions towards SpeechAce

Figure 5 revealed that students generally viewed SpeechAce positively. Nearly all found it easy to use, and the majority agreed that it helped improve their pronunciation. SpeechAce also aided students in reading phonemic transcriptions, with more than half acknowledging its effectiveness in this area. The tool was particularly helpful in addressing errors related to ending sounds, and students appreciated the detailed analysis of their pronunciation. Many students reported that they could effectively imitate the model pronunciations provided by SpeechAce and were able to learn independently using the tool.

The final open-ended question ("do you like learning pronunciation with SpeechAce? Why? Why not?) revealed that 100% of the participants (n = 34) reported enjoying learning pronunciation with SpeechAce. Notably, 76% (n = 26) expressed a strong interest in incorporating SpeechAce into all their pronunciation lessons. The most frequently cited reasons were that the tool is "interesting" (68%), "useful" (82%), and "easy to use" (74%)**.** Additionally, 88% of the participants appreciated the detailed feedback provided by SpeechAce, which helped them clearly identify specific areas of improvement. Overall, all participants (100%) agreed that SpeechAce significantly contributed to enhancing their pronunciation skills.

Results from interviews

Students' Background and Difficulties in Learning Pronunciation

When being asked about how they have been learning pronunciation previously, and whether they find them effective, the students shared similar answers. Most participants shared that they primarily learned pronunciation through formal instruction at school or English language centers. One student (S5) noted, *"I was lucky to have a native speaker teach our class for a short course—it really helped me hear the correct sounds more clearly."* Others mentioned relying on digital platforms. As one student (S7) put it, *"I often watch short clips on YouTube or Facebook—especially the ones that focus on how to pronounce difficult words."* Another (S9) added, *"Watching movies and listening to English songs has helped me a lot with pronunciation, especially with stress and intonation."* Additionally, several students mentioned using mobile apps. For example, S4 said, *"I use the Cake app almost every day because it gives me short, useful dialogues to practice with."* Similarly, S10 shared, *"TFLAT is easy to use and I like how it lets me listen and repeat after native speakers."*. However, when it comes to the effectiveness of their pronunciation learning process, most students expressed uncertainty about the improvement of their pronunciation. S9 said that "it only has short-term effectiveness. However, I can't recall what I learned later on". Similarly, S5 deemed her pronunciation learning effective, but she does not think it has been improved. S3 stated that "the pronunciation teaching methods at school may work for students with aptitude, but not for me". In terms of the difficulties they have in learning pronunciation, there were four main ones. Four students reported that different methods of teaching pronunciation from different teachers confuse them.

"there were various teachers with different accents and pronunciations during my learning process. As a result, each year we had to get used to different methods of pronunciation teaching. This, consequently caused us great confusion." – S2. This opinion is also shared by S1, S4, and S3

Five out of ten interviewees said that they do not know if their pronunciation is correct since no one helps them pinpoint their mistakes. S6 stated that "I don't know how to pronounce the word, so there should be something [like an app] or someone to correct my pronunciation", or S5 admitted that she often pronounced incorrectly, but she didn't receive any feedback or correction. S9 said that the app Cake she is currently using only grades her pronunciation on a "good", "not good enough" scale, and it doesn't offer more thorough analysis of her pronunciation. S10, a student who uses TFLAT, also shared a similar response.

Linking and ending sounds are also English pronunciation features which are troublesome to the students. S1, S7, and S8 confessed that they have trouble in pronouncing and recognizing these features in English pronunciation. S1 said she couldn't link the sound together, while S7 said she was unable to recognize fast continuous linking speech. S8 stated that she couldn't manage to pronounce the weak sounds or ending sound /t/. Having no or little knowledge in interpreting phonemic transcription is another difficulty reported by two of the interviewees, S1 and S7.

It can be inferred that the pronunciation learning of the students took place in a relatively spontaneous manner, without any fixed methods applied to their learning process. This, thus,

has caused confusion and difficulties in the way the students perceive their pronunciation competency after some time of learning it.

Students' Perceptions of Using AI to Learn Pronunciation

The data provides insights into their familiarity and perceptions regarding the use of AI in learning pronunciation. While all ten interviewees acknowledged their awareness of integrating AI in pronunciation training, their experiences and opinions varied, reflecting a complex picture. For instance, participant S1 recognized the potential of AI to enhance learning by making pronunciation adjustments more native-like. S4 highlighted the technical capabilities of AI, such as phonemic transcription. However, there was a common sentiment among students that current AI applications, including TFlat, Cake, Duolingo, and Elsa Speak, though useful, were limited in their feedback specificity. As S8 noted, these applications often provide a general percentage score of pronunciation correctness but lack detailed feedback on specific sounds or words that need improvement. This general satisfaction with AI's basic functionalities, coupled with a desire for more in-depth phonemic analysis, suggests a gap between user expectations and current AI capabilities in pronunciation training. Such findings echo the research by Chapelle (2001), who emphasized the need for AI language learning tools to provide more granular feedback to be genuinely effective. Therefore, while students are optimistic about the role of AI in pronunciation learning, there is a clear demand for advanced features that offer detailed, personalized feedback to address specific areas of pronunciation

As reported by the participants, all of them have heard and used A.I. in practicing their pronunciation at some point. However, they doubt that the A.I. they had used help them improve their pronunciation competency. They also expected the type of A.I. which further analyzes their utterance in phonemic level.

Students' Perceptions Towards Learning Pronunciation with SpeechAce

After using SpeechAce for ten weeks, the students shared their thoughts on the tool. Generally, they appreciated the detailed feedback provided by SpeechAce. S6 stated, "*I like that SpeechAce analyzes my pronunciation error in a very detailed manner, it points out what sound I got right, what I need to improve on. Furthermore, I can listen and imitate the model utterance it offers*." S4 enjoyed the ability to compare his pronunciation with the AI model, feeling that this feature significantly improved his skills. S3 highlighted that matching her pronunciation with the model often required multiple attempts, which she found both challenging and motivating.

However, some students faced difficulties while using SpeechAce. Three students cited slow internet connections as a hindrance, leading to delays or inaccurate feedback. S6 and S1 wished for an option to slow down the speech rate of the model utterances, with S1 noting, "*I know that machine [AI] doesn't guarantee 100% accuracy, but I wish there would be a function which allows you to slow the speech rate down so we can hear the utterance more clearly*." S2 found the process of accessing SpeechAce through the MLearning system inconvenient, while S7 was frustrated by inconsistent scores. Despite these issues, S4 and S8 reported no significant difficulties, apart from needing a quiet environment for practice.

When asked to compare SpeechAce with a real teacher, opinions were mixed. Four students felt that SpeechAce was comparable to a real teacher, with S9 noting, "*SpeechAce is relatively*

good, and it pronounces quite similarly to real teachers." However, others preferred learning from real teachers. S1 argued that real teachers help improve not only pronunciation but also confidence and communication skills, stating, "*Frankly, I still prefer learning with a real teacher because we can genuinely express ourselves. As a result, when we meet foreigners, we don't have the feeling of shyness. AI, however, merely helps us improve our pronunciation, not our mental confidence. That's why I think learning with real teachers is better*."

Overall, all students agreed that their pronunciation had improved after using SpeechAce. S4 said, "*I felt that my pronunciation has clearly improved since using SpeechAce*." S5 rated her improvement from a 5 or 6 before using SpeechAce to a 7 or 8 afterward. The detailed analysis and ability to retry pronunciation exercises were cited as key factors in their progress. S1 summed it up by saying, "*SpeechAce has helped me understand my pronunciation mistakes and the stress patterns, so I know where to stress and where not to*."

Discussion

The data presented previously is further discussed and synthesized with previous studies mentioned in literature review section. It aims to see how thoroughly the collected data answers the two research questions. Thus, it is divided into two parts. The first part is based on the collected data from pretest and posttest to discuss students' improvement on pronunciation competency after using SpeechAce. The second part combines data collected from questionnaire and interviews with the experimental group to conclude on students' perceptions towards using SpeechAce to practice their pronunciation.

The Impact of SpeechAce on Students' Pronunciation Competency

The pretest results show that EG's mean score was slightly higher than that of CG's mean score with 31.47 and 30.93 respectively (table 4.1). However, the Sig. (2-tailed) value obtained from the independent samples test was .671 (> .05) (table 4.2). Thus, it can be concluded that the differences between the means of CG and EG were relatively insignificant. After eleven weeks of applying SpeechAce to the EG, the posttest results show that EG's mean score is now relatively higher than CG's mean score, at 58.03 compared to 51.80 (table 4.3). Moreover, the 2-tailed Sig. value was found to be .00 (table 4.4), which is below the threshold of .05. This result indicates a statistically significant difference between the posttest means of the control group (CG) and the experimental group (EG). Therefore, it can be confirmed that the pronunciation competency of both groups improved after the treatment, with the performance of the experimental group notably surpassing that of the control group. This finding parallels with the results from the three studies previously mentioned in the literature review section by Aiello and Mongibello (2019), Moxon (2021), and Lidia (2021). However, Aiello and Mongibello acknowledged the lack of control group in their study, and Moxon He admitted the study's restrictions stemming from the absence of male students and limitations in terms of time. This study, which was done with both control group and experimental group during a process of 10 weeks, has bridged the gaps mentioned by previous authors. All these factors validated the impact of improving students' pronunciation competency through the utilizing of SpeechAce.

Students' Perceptions of Learning Pronunciation with SpeechAce

Prior to the study, it was anticipated that students might undervalue the importance of learning pronunciation, as some exhibited unintelligible pronunciation and showed limited effort in improving this skill. This perception was further influenced by the researcher's teaching experience, which involved only a few isolated sessions on pronunciation, with most instruction occurring spontaneously and informally. However, the data collected from the questionnaire revealed unexpected results. The responses indicated that all participants held a serious attitude toward learning pronunciation. Specifically, all students either agreed or strongly agreed with the statement that 'it is very important to possess good pronunciation,' and the majority expressed interest in learning pronunciation. This positive attitude was contrary to the researcher's initial expectations. Additionally, most students displayed optimism about their ability to achieve a native-like accent, aligning with the findings of Dao (2018).

Unfortunately, such enthusiasm and positivity in learning pronunciation of the students were hindered by the difficulties they had to face. In fact, many students' difficulties matched with previous points mentioned in the literature review chapter. Specifically, nearly two-thirds of the surveyed participants admitted that Students reported a lack of knowledge about effective methods for independently improving their pronunciation, despite many disagreeing or strongly disagreeing with the statement that 'teachers spend little time teaching pronunciation.' This discrepancy may be attributed to the tendency of many instructors to approach pronunciation instruction in an unstructured manner, often addressing issues reactively based on immediate classroom needs rather than through preplanned methods.(Nguyen & Newton, 2020). The challenge of lacking effective methods for learning pronunciation was also highlighted by four participants during the interviews (S1, S2, S3, S4).

> *There were various teachers with different accents and pronunciation during my learning process. As a result, each year we had to get used to different methods of pronunciation teaching. This, consequently caused us great confusion." – S2.* This opinion is also shared by S1, S4, and S3

In the previous chapter, the researcher assumed that most Vietnamese English learners had difficulty recognizing and pronouncing ending and linking sounds. Certain English phonetic sounds, including /t/, /ʧ/, /ʤ/, /ʒ/, /s/, /θ/, and /ʃ/, were identified as particularly challenging (Tam, 2005). These findings are supported by the questionnaire results: 72% of students agreed or strongly agreed that they struggled with ending and linking sounds, while 18% remained neutral and 10% disagreed. Similarly, interview responses from three participants (S1, S7, S8) echoed these concerns. In addition, a significant portion of the students (65%) reported unfamiliarity or lack of confidence in using phonemic transcriptions for learning pronunciation, while 22% expressed neutrality and 13% felt confident using them.

With the above attitude and difficulties in learning pronunciation, the students received SpeechAce very positively. First, 31 out of 32 surveyed participants said that SpeechAce was easy to use and 25 of them believed they could learn pronunciation effectively by themselves, the rest took a neutral stance. This may solve the problem of times constraint in teaching and learning pronunciation in class. Students' problem in understanding phonemic transcription, to

an extent, might be solved with SpeechAce, with nearly two thirds of the participants believing that they could read phonemic transcription, while 11 others chose neutral for this point. A similar pattern is observed in the potential of SpeechAce to assist students in addressing issues with missing ending sounds. One specific aspect of SpeechAce that students found particularly engaging was its ability to provide detailed feedback on their pronunciation. Almost all students, except one chose neutral, stated that they knew exactly which sound they pronounce incorrectly and 20 of them said they could recreate the sound by listening to and imitating the model utterance. This assertion was later developed in the interview, in which all ten interviewees reaffirmed this good point of SpeechAce. In the literature review, a number of authors mentioned the effectiveness of corrective feedback in teaching and learning pronunciation (Coupe, 2015; Ellis, 2009; Kim & Han, 2007; Lee et al., 2015; Long, 2015; Lyster et al., 2013; Saito & Lyster, 2012). Undoubtedly, SpeechAce has offered the appropriate form of corrective feedback that learners require to enhance their pronunciation learning with greater effectiveness. Finally, most students admitted that their pronunciation had improved after using SpeechAce, all participants in the interview also supported this view and half of the surveyed participants expressed that SpeechAce could help them pronounce like native speakers.

It seems that SpeechAce has broken down all barriers that stop students from learning pronunciation effectively. Also, it has helped them improve their pronunciation in certain areas. However, it should be noted that there exist some uncertainty amongs students' answers. Thus, it would never be a good idea to rely solely on SpeechAce to teach pronunciation, as many of the interviewed students reported that they still preferred real teachers to teach them pronunciation because it is not only the proper pronunciation they are looking for, they also wish to talk with real people to boost their confidence in communication. All in all, it would be wise to use SpeechAce only as an effective tool to help teachers bridge the gaps in teaching pronunciation to students, and the idea of A.I that can completely replace a human teacher has yet to come.

Conclusion

The study aims to investigate the impact of SpeechAce in developing pronunciation competency for non-English-majored students at a private university. It seeks to answer the two questions: "to what extent does SpeechAce improve students' pronunciation competency?" and "what are the students' perceptions toward learning pronunciation with SpeechAce?". A mixed method quasi-experimental design with one control group (CG) and one experimental group (EG) was applied. The results from the pretest and posttest have shown that students' pronunciation competency of the EG was improved compared to that of CG and thus, confirmed the effectiveness of using SpeechAce to improve pronunciation. The data obtained from the questionnaire and interviews have also revealed that the majority of students received SpeechAce positively although there lie some uncertainties which calls for teachers' careful consideration and preparation when using SpeechAce to teach pronunciation.

For educators, the study underscores the importance of recognizing students' enthusiasm for learning pronunciation, aligning with Gilbert's (2008) emphasis on leveraging students' willingness to improve. Despite the challenges, such as the time-consuming nature of pronunciation instruction, teachers have a critical role in making pronunciation learning meaningful and engaging for students. Derwing and Munro (2005) highlight the necessity for teachers to not only instruct but also model accurate pronunciation to facilitate students' acquisition. Additionally, teachers must focus on correcting common pronunciation errors specific to Vietnamese English learners, as indicated by Jenkins (2000). Incorporating phonemic transcription into teaching, as suggested by Celce-Murcia et al. (2010), is crucial, as it has been proven in this study to significantly benefit students' learning process. Above all, embracing AI technology, such as SpeechAce, and integrating it into teaching practices can enhance the effectiveness of both learning and teaching, as recommended by Chapelle (2001). For students, SpeechAce provides a valuable resource for independent pronunciation practice. However, effective use of this tool requires students to become proficient in interpreting phonemic transcriptions, a skill essential for pronunciation learning, as noted by Levis (2005). This also necessitates a degree of self-discipline for consistent practice.

For curriculum designers, the positive impact of SpeechAce underscores the need for incorporating AI technology into English language education, supporting Pennington's (1996) argument for technology integration in language learning. Additionally, formal pronunciation training for teachers can enhance their confidence and competence, thereby benefiting students, as emphasized by Derwing and Munro (2015). Institutions could integrate SpeechAce into curricula through blended learning models, where students receive immediate, individualized pronunciation feedback alongside traditional instruction

While the study achieved its objectives, certain limitations must be acknowledged. The study's sample size was constrained by practical considerations, potentially limiting the generalizability of the findings to broader populations. Additionally, the study focused on a specific demographic, and the impact of SpeechAce on learners from diverse linguistic backgrounds might differ. The participants' familiarity with technology, particularly ASR-based tools like SpeechAce, could have influenced their engagement and outcomes, and the transformation of textbook exercises into the SpeechAce format was not always accurate, potentially affecting the validity of the activities. Furthermore, the pretest and post-test design did not involve oral pronunciation tasks, which could have provided a more reliable assessment of pronunciation competency. Finally, the lack of inter raters is another factor that does not guarantee a high level the reliability and validity of the study.

Future research should involve larger sample sizes to assess the effectiveness of ASR technology in teaching pronunciation. Pretests and post-tests should be designed to evaluate all components of English pronunciation and include oral tasks to accurately gauge competency. The inclusion of inter-raters in grading will enhance the objectivity of results. Additionally, further studies should explore the most effective pedagogical strategies for integrating ASR into language instruction, considering diverse learner populations and investigating the impact of ASR on other language skills, such as listening comprehension and speaking fluency.

References

Bajorek, J. (2017). L2 pronunciation tools: The unrealized potential of prominent computer-assisted language learning software. *Issues and Trends in Educational Technology*, *5*(1). https://doi.org/10.2458/azu_itet_v5i1_bajorek

Celce-Murcia, M., Brinton, D., & Goodwin, J. (2010). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, *30*(2). https://doi.org/10.1016/S0346-251X(02)00012-X

Fraser, H. (2000). Coordinating improvements in pronunciation teaching for adult learners of English as a second language. Canberra: DETYA (ANTA Innovative Project). Retrieved from https://helenfraser.com.au/wp-content/uploads/ANTA-REPORT-FINAL.pdf

Gilakjani, A. P. (2012). The significance of pronunciation in English language teaching. *English Language Teaching*, *5*(4), 96-107. https://doi.org/10.5539/elt.v5n4p96

Gorjian, B., Hayati, A., & Pourkhoni, P. (2013). Using Praat software in teaching prosodic features to EFL learners. *Procedia - Social and Behavioral Sciences*, *84*. https://doi.org/10.1016/j.sbspro.2013.06.505

Ha, T. T. X., & Bao, N. K. V. (2023). Common errors in pronunciation of non-English majored students at the university of Transport and Communication Ho Chi Minh Campus. *South Asian Research Journal of Arts, Language and Literature*, *5*(03), 64-70. https://doi.org/10.36346/sarjall.2023.v05i03.002

Hassim, H., Kassim, H., Kassim, A., & Kassim, M. (2023). Exploring the use of artificial intelligence-based technology to enhance creativity in ESL speaking classroom. *The Proceedings of 7th IEEE Congress on Information Science and Technology (CiSt)*. https://doi.org/10.1109/CiSt56084.2023.10409965

Hirai, A., Kovalyova, A. (2023). Using speech-to-text applications for assessing English language learners' pronunciation: A comparison with human raters. In Suárez, MdM., El-Henawy, W.M. (eds) *Optimizing Online English Language Learning and Teaching. English Language Education*, vol 31, 337-355. Springer, Cham. https://doi.org/10.1007/978-3-031-27825-9_17

Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112-126. https://doi.org/10.1017/s0267190501000071

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2023). Effects of automatic speech recognition technology on EFL learners' willingness to communicate and interactional features. *Educational Technology and Society*, *26*(3), 37-52. https://doi.org/10.30191/ETS.202307_26(3).0004

Kholis, A. (2021). Elsa speak app: Automatic speech recognition (ASR) for supplementing English pronunciation skills. *Pedagogy: Journal of English Language Teaching*, *9*(1), 1-14. https://doi.org/10.32332/joelt.v9i1.2723

Kumar, L. A., & Renuka, D. K. (2023). *Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision*. CRC Press. https://doi.org/10.1201/9781003348689-5

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, *27*, 184-202. https://doi.org/10.1017/S0267190508070098

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*(3), 369-377. https://doi.org/10.2307/3588485

Li, J., Deng, L., Haeb-Umbach, R., & Gong, Y. (2015). *Robust Automatic Speech Recognition: A Bridge to Practical Applications*. eBook ISBN: 9780128026168

Loc, N. T., & Newton, J. (2020). Pronunciation teaching in Tertiary EFL classes: Vietnamese teachers' Beliefs and Practices. *The Electronic Journal for English as a Second Language, 24*(1), 1-20.

Moxon, S. (2024) *A review of the website pronounce.com*, *Computer-Assisted Language Learning Electronic Journal*. Available at: https://callej.org/index.php/journal/article/view/22(Accessed: 09 May 2025).

Moxon, S. (2021). Exploring the effects of automated pronunciation evaluation on L2 students in Thailand. *IAFOR Journal of Education*, *9*(3), 41-56. https://doi.org/10.22492/ije.9.3.03

Nguyen, H. N. (2023). Teaching English sounds to Vietnamese secondary school students: From theories to applications using learner-centered methods. *International Journal of Language Instruction*, *2*(1), 16-29. https://doi.org/10.54855/ijli.23212

Vu, N. N. (2024). *Computational Linguistics: From Theory to Practice*. Hồ Chí Minh: NXB ĐHSP TP.HCM.

Nguyen, T. (2017). Vietnam's national foreign language 2020 project after 9 years: A difficult stage. *The Asian Conference on Education and International Development*.

Sholekhah, M. F., & Fakhrurriana, R. (2023). The use of ELSA Speak as a mobile-assisted language learning (MALL) towards EFL students' pronunciation. *JELITA: Journal of Education, Language Innovation, and Applied Linguistics*, *2*(2), 93-100. https://doi.org/10.37058/jelita.v2i2.7596

Spring, R. and Tabuchi, R. (2024) *The role of ASR training in EFL pronunciation improvement: An in-depth look at the impact of treatment length and guided practice on specific pronunciation points*, *Computer-Assisted Language Learning Electronic Journal*. Available at: https://callej.org/index.php/journal/article/view/417 (Accessed: 09 May 2025).

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism*, *15*(4), 905-916. https://doi.org/10.1017/S1366728912000168

Vu, N. N. (2016a). An investigation of Vietnamese students' learning styles in online language learning. *Journal of Science, Ho Chi Minh University of Education*, *79*(1), 16-24.

Vu, N. N. (2016b). Mobile learning in language teaching context of Vietnam: An evaluation of students' readiness. *Journal of Science, HCMC University of Education*, *7*(85), 16-27. Retrieved from http://www.vjol.info/index.php/sphcm/article/viewFile/24861/21273

Biodata

Assoc. Prof. Dr. Nguyen, Ngoc Vu is an Associate Professor of Linguistics at Ho Chi Minh City University of Foreign Languages - Information Technology (HUFLIT), Vietnam and Chairman of STESOL, a TESOL organization setup by Association of Vietnamese Universities and Colleges. His research interests include Computational Linguistics, Teaching English to Speakers of Other Languages (TESOL) and Computer-Assisted Language Learning (CALL).

Mr. Vo, Thiet Thach holds a Master of Arts in TESOL and is an experienced English lecturer at Ho Chi Minh City University of Foreign Languages - Information Technology, Vietnam. With over 10 years of university-level teaching, his research interests include English for Specific Purposes (ESP), Computer-Assisted Language Learning (CALL), and Linguistics.

Dr. Tran, Vu Diem Thuy is currently working as Head of Department of Applied Linguistics at the Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages - Information Technology, Vietnam. She has been teaching English for over 22 years at the university level. Her research interests include ESP, CALL/MALL, Linguistics, Computational Linguistics, and Teacher Education.