



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 8, 103599 (2026)

<https://doi.org/10.29140/tltl.2026.103341>

Segmental and Suprasegmental Pronunciation Training via AI: An Exploration of University Students' Perceptions and Attitudes



NGUYEN HUU LAM KHANH^a 

TRAN THANH TAN (corresponding author)^b 

^a*Ho Chi Minh City Open University, Vietnam*
lamkhanh.350471@gmail.com

^b*Ho Chi Minh City Open University, Vietnam*
tan.tt@ou.edu.vn

Abstract

Artificial Intelligence (AI) has emerged as a transformative tool in language learning, particularly for improving pronunciation. Although recent studies have highlighted AI's effectiveness in pronunciation training, there has been a limited in-depth investigation into the perceptions of university-level English majors regarding the use of AI for pronunciation practice, especially focusing on two key features: segmental and suprasegmental aspects. Therefore, this research aimed to explore the perceived benefits, challenges, and attitudes of English as a Foreign Language (EFL) students, specifically university-level English majors, using AI to practice these two pronunciation features. A mixed-methods design was employed, combining a questionnaire with semi-structured interviews and journal diaries. The researchers constructed a questionnaire focusing on segmental and suprasegmental features among 70 participants in a university in Vietnam recruited by purposive sampling. This questionnaire was constructed by the researcher with critical reviews from six experts in the Teaching English to Speakers of Other Languages (TESOL) field. Qualitative data from semi-structured interviews and journal diaries were collected from five participants recruited by convenience sampling. The qualitative instruments were theoretically based on the Tripartite Model of Attitudes (Rosenberg & Hovland, 1960). Quantitative data were analyzed in SPSS 26 using Cronbach's alpha and descriptive statistics, while qualitative data were coded through deductive thematic analysis to generate themes. Findings revealed that students found AI tools helpful for improving segmental pronunciation features, especially long vowels, but encountered challenges such as technical inaccuracies. Students showed moderate improvement in their suprasegmental features, like word stress, while receiving insufficient

Copyright: © 2026 Nguyen Huu Lam Khanh & Tran Thanh Tan. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. **Data Availability Statement:** All relevant data are within this paper.

feedback for pauses, speech chunking, and connected speech. Despite these challenges, most students expressed enthusiasm for AI's convenience and personalized support. This study presents implications for theory and offers practical insights for teachers, English majors, and AI developers.

Keywords: AI in pronunciation; segmental features; suprasegmental features; EFL university students; mixed-methods research

Introduction

In today's globalized society, English has emerged as a universal medium of communication across various sectors, including education, business, and international relations. Among the multiple skills that constitute English proficiency, pronunciation holds a particularly vital role in ensuring effective communication and academic success (Jenkins, 2000). As Jenkins (2000) and Brown (2014) emphasized, accurate pronunciation not only supports intelligibility in conversation but is also crucial for high performance in standardized English tests such as International English Language Testing System (IELTS) and Test of English as a Foreign Language (TOEFL), which comprehensively assess speaking.

However, mastering English pronunciation is often one of the most significant challenges for English as a Foreign Language (EFL) learners due to its complexity. According to Gilakjani (2012, 2016), learners frequently struggle with the pronunciation of English words, a problem largely attributed to the contrast between the phonological systems of their native and target languages. Pronunciation comprises two key dimensions: segmental features, such as vowels and consonants, and suprasegmental features, including intonation, stress, and rhythm (Celce-Murcia et al., 2010; Ladefoged & Johnson, 2011). These elements jointly contribute to a speaker's comprehensibility, and difficulties in either can result in communication breakdowns (Yates & Zielinski, 2009).

To overcome these challenges, many learners are turning to technology-based interventions, particularly Artificial Intelligence (AI) (Vančová, 2023). In the field of language acquisition, AI has emerged as a promising tool, offering personalized and intelligent support. As Vančová (2023) noted, AI-driven tools like chatbots and pronunciation apps can analyze speech patterns, diagnose specific errors, and deliver individualized feedback. These capabilities have demonstrated significant potential for improving learners' pronunciation. Research by Junaidi (2020) indicated that students using AI tools showed superior speaking skills compared to those who were taught without modern technology. Dennis (2024) and Selladurai and Pandey (2024) further reinforced these findings, noting that AI applications such as English Language Speech Assistant (ELSA) Speak significantly enhance learners' pronunciation accuracy through personalized instruction and real-time corrections.

Despite the encouraging evidence, a comprehensive understanding of how learners, particularly English majors, perceive the benefits and limitations of using AI for pronunciation practice remains lacking. Additionally, there is insufficient insight into their attitudes toward using AI tools to address both segmental and suprasegmental aspects of pronunciation. This study, therefore, aims to fill this gap by exploring students' perceptions and attitudes towards AI-supported pronunciation practice. Investigating students' perceptions and attitudes toward AI-assisted pronunciation practice is essential because their perceptions shape how they interpret the usefulness and effectiveness of AI tools, while their attitudes, comprising emotional, behavioral, and cognitive dimensions, influence their engagement and overall language learning success (Rosenberg & Hovland, 1960; Ajzen, 2001).

In terms of practical aspects, English majors face ongoing difficulties in producing accurate pronunciation (Phi & Mai, 2024), especially in managing both segmental and suprasegmental features. These pronunciation deficiencies can lead to poor oral performance, reduced confidence, and limited communicative competence (Bakar & Abdullah, 2015; Kobilova, 2022). Such challenges can negatively impact students' academic performance and professional opportunities, as miscommunication may lead to misunderstandings and missed opportunities for advancement (Brown, 2014).

From a theoretical perspective, while recent studies have explored the use of AI tools in pronunciation training, gaps remain as most studies assess pronunciation broadly without distinguishing segmental and suprasegmental features or focus on only one aspect, limiting comprehensive insights into pronunciation development (Anggraini, 2022; Hoang et al., 2023; Nguyen et al., 2025), particularly at the university level and within the context of English majors. Additionally, many existing studies treat pronunciation as a general skill, without analyzing learners' responses to training in the specific domains of segmental and suprasegmental features. Furthermore, while global literature on AI in education is growing, studies in the Vietnamese context remain limited.

This study seeks to fill these gaps with two primary aims. First, it seeks to investigate the perceived benefits and challenges that English majors face when using AI-based tools to practice both segmental and suprasegmental aspects of pronunciation. Second, it aims to examine students' attitudes towards the use of AI in pronunciation training, how they feel about, interact with, and evaluate the role of such technology in their learning process. These aims are expected to generate insights into the practical application of AI tools and their psychological and pedagogical impacts on learners.

Practically, this study provides valuable insights into how English majors perceive and utilize AI tools to enhance their pronunciation. Such understanding can guide educators in integrating AI into language instruction more effectively, leading to improved pronunciation training outcomes. Furthermore, it may inform the design of AI tools that are more responsive to students' specific pronunciation needs.

Theoretically, the research contributes to the relatively underexplored area of AI application in pronunciation teaching. While AI has been used in vocabulary and grammar instruction, fewer studies have delved into its impact on segmental and suprasegmental pronunciation. By exploring learners' perceptions and attitudes in these two dimensions, this study extends the body of knowledge on AI-enhanced language learning. It also offers recommendations on how AI technologies can be optimized for pronunciation instruction, thereby contributing to future tool development and pedagogical practices.

RQ1. What are the perceptions of English major students regarding the benefits and challenges of using AI tools to practice segmental and suprasegmental features of pronunciation?

RQ2. How do they describe their attitudes towards using AI tools to practice segmental and suprasegmental features of pronunciation?

Literature Review

Definitions and Components of English Pronunciation

A comprehensive understanding of pronunciation in language learning begins with its definition and components. Pronunciation is broadly defined as the production of sounds that convey meaning (Otlowski, 1998; Yates & Zielinski, 2009). According to Ladefoged and Johnson (2011), English pronunciation comprises segmental and suprasegmental features. Segmental features involve individual phonemes, such as vowels and consonants, whose accurate articulation is crucial for avoiding

misunderstandings (Celce-Murcia et al., 2010). In contrast, suprasegmental features, including stress, intonation, rhythm, and connected speech, contribute to conveying meaning, emotion, and nuance in spoken interaction (Zuhairya et al., 2024). Together, these components interact to shape overall speech intelligibility (Yates & Zielinski, 2009).

AI and Pronunciation in Language Learning

The significance of pronunciation lies in its role as a foundational element of effective communication. Even when learners accurately articulate individual phonemes, communication can still break down due to inappropriate stress or intonation (Levis, 2005; Kobilova, 2022). Mispronunciation can lead to miscommunication and negatively affect both academic and professional interactions; these difficulties are further exacerbated by limited exposure to authentic speech and a lack of individualized feedback (Derwing & Munro, 2015; Albiladi, 2019; Ngo et al., 2024).

Artificial Intelligence (AI) tools have emerged as a potential solution to these issues. These tools, incorporating advanced speech recognition and adaptive learning technologies, provide real-time feedback, phonetic training, and personalized instruction (Anggraini, 2022; Dennis, 2024; Lee, 2025; Sánchez et al., 2024), so that these tools are accessible and beneficial to students at all levels of English proficiency (Busso & Sanchez, 2024). AI-powered applications such as ELSA Speak and Speechling not only correct pronunciation errors instantly but also tailor learning to individual needs (Babaeian, 2023; Zhao, 2024). Compared to traditional methods, which often lack immediate feedback and individualization, AI-based systems foster learner autonomy, consistency, and engagement (Dhanapal et al., 2024).

To explore students' perceptions and attitudes toward these tools, this study adopts the Tripartite Model of Attitudes (Rosenberg & Hovland, 1960), which divides attitudes into three components: affective (feelings), behavioral (actions), and cognitive (beliefs). This model has been successfully applied in prior research to evaluate learners' engagement with AI tools in language learning, particularly with applications like Grammarly and AI-speaking assistants (Khan et al., 2024; Tran, 2024; Nguyen et al., 2024). It provides a holistic framework for assessing students' emotional responses, behavioral tendencies, and cognitive judgments about AI-assisted learning, thereby forming the backbone of this study's analytical approach.

Previous Studies

Research on AI-supported pronunciation tools has grown significantly in recent years and can be grouped both globally and within the Vietnamese context. Globally, Anggraini (2022) conducted classroom action research using ELSA Speak with 30 basic-level learners. The study observed consistent improvements in learners' pronunciation over three cycles and reported positive perceptions regarding ELSA's usability and feedback. Similarly, Marlinda and Huda (2024) applied a quasi-experimental design with senior high school students and found that the experimental group using ELSA Speak achieved significantly better post-test scores than the control group. Learners also showed greater self-confidence and speaking proficiency. Dennis (2024) used a mixed-methods design with intermediate-level university students to assess Speechling's effectiveness. Results confirmed notable improvements in pronunciation and speaking skills, with students praising the system's speech recognition and feedback features. These studies collectively affirm the value of AI tools in enhancing learner pronunciation across various age and proficiency levels.

Within Vietnam, several recent studies have highlighted the positive impact of AI-assisted tools on pronunciation training. Hoang et al. (2023) found that using MissionFluent significantly improved stress and intonation among vocational students. Vu (2024) and Nguyen et al. (2025) both reported notable

gains in pronunciation through ELSA Speak, with students praising its real-time feedback and personalized learning, despite some technical limitations. Similarly, Nguyen and Nguyen (2024) observed generally positive attitudes toward ELSA Speak among non-English majors at Can Tho University. Nguyen et al. (2025) demonstrated that SpeechAce enhanced pronunciation scores in a quasi-experimental setting, with students valuing its usability but still preferring teacher guidance. Collectively, these studies affirm that AI tools like ELSA Speak and SpeechAce effectively support pronunciation improvement and learner engagement through instant, adaptive feedback.

Identification of Research Gaps

Despite the growing body of research supporting the use of AI in pronunciation training, several theoretical and methodological gaps remain. Theoretically, there have been some gaps. Firstly, most studies have evaluated pronunciation as a general skill without distinguishing its core components, segmental and suprasegmental features, or focused on only one feature when assessing learners' perceptions, limiting the comprehensiveness of findings on pronunciation development (Anggraini, 2022; Hoang et al., 2023; Dennis, 2024; Marlinda & Huda, 2024; Nguyen & Nguyen, 2024; Vu, 2024; Nguyen et al., 2025). Very few investigations have examined how AI tools specifically address these two dimensions. Consequently, the comparative effectiveness of AI tools in tackling segmental issues like consonant/vowel errors versus suprasegmental elements such as intonation and rhythm remains unclear.

Secondly, there is a lack of research focused on university-level English majors, a group for whom accurate pronunciation is especially crucial for academic and professional development. Most existing studies have centered on high school students or non-English majors, thereby overlooking the perspectives of a key demographic. Thirdly, several studies focus on learners' perceptions of a single AI-powered tool without comparing multiple platforms or addressing how different tools support varied pronunciation features (Nguyen et al., 2025). Fourthly, the application of robust theoretical frameworks to evaluate learner attitudes is scarce. The Tripartite Model of Attitudes, despite its relevance, has rarely been employed in this context. As a result, existing studies often lack a comprehensive analysis of learners' emotional, behavioral, and cognitive responses to AI tools.

Methodologically, a dominant reliance on quasi-experimental and pretest-posttest designs has left little room for exploring learners' subjective experiences and nuanced perceptions. Questionnaires used in previous studies often lacked theoretical grounding and expert validation, raising concerns about their content validity and the reliability of their findings (Yaghma, 2003; Nguyen et al., 2025). Additionally, interviews conducted in many studies were not conducted with theoretical frameworks, leading to data that may not fully capture learners' complex interactions with AI tools (Anggraini, 2022; Dennis, 2024; Marlinda & Huda, 2024). In addition, some studies remain limited in sample size, which restricts the generalizability of findings (Nguyen et al., 2025).

To address these gaps, the present study aims to provide a more detailed exploration into English majors' practice of segmental and suprasegmental features of pronunciation. It applies the Tripartite Model of Attitudes to capture a nuanced understanding of students' emotional reactions, learning behaviors, and cognitive evaluations. Methodologically, it adopts a mixed-methods approach combining expert-validated questionnaires and structured interviews with a larger scale of participants to ensure content validity and data richness.

Theoretical Framework: A Tripartite Perspective

To explore students' attitudes toward these tools, this study adopts the Tripartite Model of Attitudes (Rosenberg & Hovland, 1960), which divides attitudes into three components: affective (feelings),

behavioral (actions), and cognitive (beliefs). The affective aspect refers to an individual's emotional response toward an object or situation, which influences motivation.

The behavioral component reflects how attitudes translate into actions; for instance, positive attitudes toward learning foster greater classroom participation. Meanwhile, the cognitive element involves one's beliefs and thoughts about the attitude's object.

This framework has been successfully applied in prior research to evaluate learners' engagement with AI tools in language learning, particularly in applications such as Grammarly and AI-speaking assistants (Khan et al., 2024; Tran, 2024; Nguyen et al., 2024). It provides a holistic framework for assessing students' emotional responses, behavioral tendencies, and cognitive judgments about AI-assisted learning, thereby forming the backbone of this study's analytical approach.

Methodology

Research Design

This study adopts an embedded mixed-methods design, a research approach where quantitative and qualitative data are collected simultaneously, with qualitative data embedded within the quantitative framework (Creswell & Plano Clark, 2017). This research design ensures that the study will capture both qualitative and quantitative understanding for a well-rounded knowledge of the research topic. In this research, the primary data source was a quantitative questionnaire that assessed perceptions toward using AI in pronunciation practice. Meanwhile, the qualitative data from interviews and journals enriched the interpretation by capturing perceptions and attitudes, thereby offering a more comprehensive understanding of the phenomenon.

Participants and Sampling

The sample for the quantitative phase consisted of 70 second-year English-majoring students at a public university in the South of Vietnam. They were selected through purposive sampling, which is appropriate when participants are chosen based on specific, study-relevant characteristics (Fraenkel et al., 2011). This group, including eight males and 62 females, practiced pronunciation using two free-version AI-powered websites, as illustrated with Figure 1 and Figure 2 (SpeechAce and IELTSAce), for over six weeks.

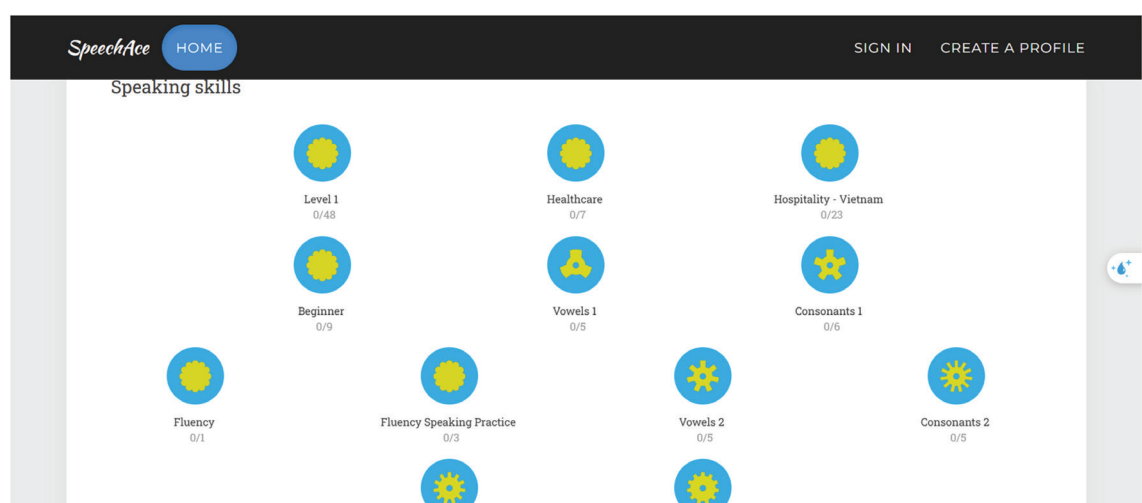


Figure 1 A screenshot to Illustrate SpeechAce Website



Figure 2 A Screenshot to Illustrate IELTSAce Website

During the six-week practice, students were asked to use at least two of the three AI-powered websites per session, selecting freely based on their preference. They practiced three to four times a week for about ten minutes each time, focusing on both segmental and suprasegmental pronunciation through tasks like reading aloud, free speaking, and responding to questions.

For the qualitative phase, five female participants were selected from this group using convenience sampling that involves selecting individuals based on their accessibility and willingness to join (Creswell & Guetterman, 2019). These participants volunteered for semi-structured interviews and journal diaries, offering in-depth reflections on their practice process with the AI tools.

Data Collection Methods

This study employed three main instruments: questionnaires, interview protocols, and journal diary templates, creating data triangulation. First, the questionnaire comprised 12 items, divided into two sections: demographic information and participants' perceptions of the AI tools, specifically regarding segmental and suprasegmental pronunciation. It was reviewed by six experts and piloted for cognitive interviews with five target participants, then administered after 70 participants completed six weeks of practicing with AI-based platforms. Hair et al. (2009) stated that the number of participants should be five times the number of the items, so 12 items of the questionnaire, with the participation of 70 students, can prove valuable in the statistics.

Secondly, the semi-structured interview protocol was designed to elicit students' personal perceptions and attitudes toward AI tools. The interviews were completely conducted in Vietnamese to help participants express their ideas more comfortably and accurately without language barriers. The questions reflected participants' perceptions of benefits and challenges when practicing pronunciation with the AI tools and the cognitive, affective, and behavioral dimensions of the Tripartite Model of Attitudes (Rosenberg & Hovland, 1960), ensuring consistency with the study's theoretical framework. Finally, complementing these instruments, participants also kept journal diaries after each practice session throughout the six-week practice phase. These diaries, guided by open-ended questions aligned with the Tripartite Model of Attitudes, encouraged students to document their thoughts, emotions, and behaviors related to AI-assisted pronunciation training. The reflective entries provided insights into how their learning experiences evolved over time.

Data Analysis Methods

Quantitative data from the questionnaires were analyzed using SPSS 26. Key analyses included: Cronbach's Alpha to assess internal consistency and reliability (Tavakol & Dennick, 2011), and descriptive statistics (mean, standard deviation) to summarize responses and reveal trends.

Qualitative data from interviews and diaries were analyzed using deductive thematic analysis (Braun & Clarke, 2006). This six-step process included familiarization with data, generating codes, identifying and reviewing themes, defining theme names, and producing final reports. This method allowed for flexible, in-depth interpretation of the participants' lived experiences.

Validity and Reliability

In terms of quantitative data, to ensure reliability, the researcher calculated a Cronbach's Alpha of 0.976, indicating excellent internal consistency across the 12 questionnaire items. The corrected item-total correlations ranged from 0.848 to 0.896, showing that all items strongly contributed to the overall reliability.

To assess content validity, the questionnaire was reviewed by six experts from Vietnam, Thailand, and the U.S. through a formal expert review process. Four experts rated the items, yielding a Content Validity Index (I-CVI) of 1, meeting the standard threshold for four raters (Polit & Beck, 2006). Additionally, cognitive interviews with five target participants were conducted to ensure item clarity. These participants read the questionnaire in both English and Vietnamese and provided feedback, which was used to revise ambiguous wording.

To assess construct validity, the researcher conducted the KMO and Bartlett's Test, yielding a KMO value of 0.934, which indicates excellent sampling adequacy (Kaiser, 1974). Bartlett's Test of Sphericity was also significant ($\chi^2 = 1164.347$, $df = 66$, $p = .000$), confirming that the data is appropriate for factor analysis (Hair et al., 2009). These results demonstrate that the questionnaire has strong construct validity.

As for the qualitative data, trustworthiness was ensured through credibility, transferability, and confirmability. Credibility was established through member checking, in which participants reviewed their interview transcripts, while the researcher maintained prolonged engagement to build trust and gain deeper insights into participants' experiences. Transferability was supported by clearly describing the sampling method and participant demographics, which allows other researchers to determine the applicability of the findings to similar educational contexts. Confirmability was achieved through data triangulation, combining information from questionnaires, interviews, and journal diaries to reduce potential researcher bias and enhance the objectivity of the findings.

Ethical Considerations

Ethical procedures were strictly followed throughout the study. Participants were fully informed about the research goals, data collection process, and their rights. Before taking part in interviews, each participant signed a consent form, confirming their voluntary participation and understanding of confidentiality procedures.

During data collection, anonymity and confidentiality were maintained. Participants' names were hidden, and all personal data were securely stored. Communication regarding the journal diary instructions and follow-ups was conducted through Zalo, a user-friendly and accessible platform for Vietnamese students. Participants were regularly reminded of their right to withdraw at any point without any consequence.

Results

Perceptions Toward the Use of AI Tools in Pronunciation Practice

Table 1 *Descriptive Statistics of Segmental Features*

	Mean	SD
1. AI tools can help me improve my pronunciation with short vowel sounds (/ɪ/, /æ/, /e/, /a/, /ʌ/, /ʊ/, /ə/)	3.94	.976
2. AI tools can help me improve my pronunciation with long vowel sounds (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/)	3.99	.955
3. AI tools can help me improve my pronunciation with diphthongs (/ɪə/, /eə/, /eɪ/, /ɔɪ/, /aɪ/, /əʊ/, /aʊ/, /ʊə/)	3.89	1.071
4. AI tools can help me improve my pronunciation with triphthongs (/aʊə/, /aɪə/, /eɪə/, /əʊə/, /ɔɪə/)	3.81	1.054
5. AI tools can help me improve my pronunciation with voiced consonant sounds (/b/, /d/, /g/, /v/, /z/, /ð/, /dʒ/, /ʒ/)	3.89	1.071
6. AI tools can help me improve my pronunciation with voiceless consonant sounds (/p/, /t/, /k/, /f/, /s/, /θ/, /tʃ/)	3.87	1.076
7. AI tools can help me improve my pronunciation with other consonant sounds (/m/, /n/, /ŋ/, /l/, /j/, /h/, /r/, /w/)	3.93	1.026

Students perceived AI-powered tools as largely beneficial in improving pronunciation, particularly in segmental features. Data from 70 questionnaire responses, as shown in Table 1, indicated generally favorable opinions across various sound categories. Long vowel sounds received the highest mean score ($M = 3.99$, $SD = 0.955$), followed by short vowels ($M = 3.94$), and other consonant sounds ($M = 3.93$). Diphthongs, voiced consonants, and voiceless consonants also received positive scores ($M \approx 3.87 - 3.89$), though with slightly higher variability in opinion. Triphthongs had the lowest score ($M = 3.81$), reflecting more uncertainty about the tools' effectiveness in this area.

Qualitative data confirmed these findings. Students appreciated the immediate corrective feedback and the ability to distinguish between similar sounds. For instance, P2 described *"I realized I often confused the short vowel in "sit" with the long vowel in "seat," and the immediate corrections improved my accuracy significantly."* while P5 noted an improvement in distinguishing between /θ/ and /s/ through practice with IELTSAce.

Despite these benefits, challenges remained. Participants criticized the feedback for being too general or difficult to interpret. P3 and P4, for example, suggested that AI should highlight specific errors more clearly rather than marking an entire sentence. P5 added that being told she read a sentence correctly, even when she knew there were minor issues, limited her ability to recognize needed improvements.

One challenge I faced is that it's easy to get confused because if I read something 70% correctly, the tool might tell me that I've read the entire sentence correctly. This could make it harder to spot areas where I still need improvement. (P5)

Technology-related concerns also surfaced. Students reported occasional issues with speech recognition, especially when speaking softly or with an accent. For instance, P4 expressed frustration at repeated misinterpretations of her pronunciation, even after multiple attempts. P5 encountered similar

problems: “AI often marked my pronunciation as wrong even though I had practiced that word many times, but it kept showing as incorrect. For example, the word “I”, I could never get it right while practicing with AI.”

Table 2. *Descriptive Statistics of Suprasegmental Features*

	Mean	SD
1. AI tools can help me improve my pronunciation of individual words with the correct stress.	3.81	1.171
2. AI tools can help me improve my pronunciation of complete sentences with correct intonation.	3.66	1.153
3. AI tools can help me improve my pronunciation in connected speech (elision, linking, assimilation, intrusion, etc.).	3.61	1.107
4. AI tools can help me improve my use of pauses and speech chunking.	3.57	1.187
5. AI tools can help me improve my rhythm in speaking English.	3.71	1.038

In terms of suprasegmental features, Table 2 illustrates that students’ perceptions were generally positive but less enthusiastic than for segmental features. Word stress was rated the highest ($M = 3.81$), followed by rhythm ($M = 3.71$) and sentence intonation ($M = 3.66$). Pauses and chunking received the lowest average rating ($M = 3.57$). Students acknowledged the tools’ support for improving rhythm and stress but expressed skepticism about the tools’ ability to enhance more nuanced elements like connected speech and intonation.

Journal and interview data supported this trend. P5 expressed “*This app helped me realize the importance of stress in words. For instance, the difference between PREsent (a noun meaning “gift”) and preSENT (a verb meaning “to show or introduce”) highlights how stress can change the meaning of a word entirely.*” while others like P4 mentioned “*I usually spoke in a flat tone. When practicing with the AI, it didn’t give explicit guidance, but when I listened to the playback, I could notice differences between my voice and the AI’s voice. From those differences, I was able to improve my intonation, especially in terms of pitch variation.*”

P1 and P3 both commented that AI tools tended to focus heavily on segmental aspects, leaving suprasegmental elements largely unaddressed. Additionally, P2 noted, “*Actually, I don’t think my intonation improved much because I rarely saw the AI give any feedback on it. Things like falling or rising tones were hardly ever mentioned in its evaluation.*”

Attitudes Toward AI Tools: Tripartite Model of Attitudes

The findings revealed that students’ attitudes toward AI pronunciation tools reflect cognitive, affective, and behavioral dimensions. Cognitively, they believed AI effectively identified errors through visual and phoneme feedback, though some questioned repetitive or shallow responses. Affectively, students felt engaged, confident, and satisfied, yet sometimes nervous or discouraged by unclear scores. Behaviorally, they practiced more frequently, repeated mistakes, and planned continued use.

Cognitive Component – Beliefs About Effectiveness

Cognitively, participants generally believed that AI tools were useful for identifying and correcting pronunciation errors. All five interviewees agreed that the tools helped highlight mistakes they might

otherwise overlook. P1 and P2 found the visual feedback and phoneme-specific correction particularly helpful, *“The AI tool can help assess my pronunciation mistakes by highlighting the words that I mispronounce”* (P1), *“It’s like, normally I wouldn’t think I pronounced that word wrong, but when practicing with the AI... I realized, oh, I’ve actually been mispronouncing it all along, so I was quite surprised”* (P2). Similarly, P5 appreciated features that allowed her to hear correct pronunciation alongside her own.

It provided a function that allowed me to see syllables and phoneme mistakes in my sentences, and it also had a feature that enabled me to hear the correct pronunciation. So, it is really convenient for practicing and remembering words. (P5)

Nevertheless, some limitations were noted. P4 expressed concerns about repetitive and superficial feedback, stating that the tools tended to focus on the same error repeatedly without addressing deeper issues. She said, *“When I practiced with the AI, sometimes I repeated a word many times, but it kept showing the same error without identifying any new ones, so I didn’t know what exactly to improve.”* Despite this, participants reported noticeable improvement in pronunciation accuracy, especially in problematic sounds like /ʃ/ and final consonants. As P5 described, *“My teacher often pointed out that I had trouble with the /f/ sound, like in the word ‘shoe.’ I didn’t pay much attention to that sound and just pronounced it like a regular ‘s.’ But since I started practicing with AI, it kept marking that word as incorrect, so I had to practice it repeatedly. Now, I remember it and have been able to correct my pronunciation.”* These improvements were credited to the repetitive nature of AI practice and the detailed visual feedback.

Affective Component – Emotional Responses

On the one hand, most participants expressed positive feelings about using AI tools. They found the experience engaging and even enjoyable. For example, P4 described:

I felt excited when I wanted to improve my midterm speaking score. At that time, I didn’t have an environment to practice, so I turned to AI for training. I was really engaged with AI because I was truly focused on it, pushing myself to come up with answers. It felt like I was fully concentrating, wanting AI to give me a specific score so that I could understand how to improve my midterm grade. (P4)

Others, such as P1 and P3, enjoyed the interactive nature of the platforms, particularly the read-aloud and free-speaking features. P1 said, *“I think the feature where the AI asks me questions directly and I have to respond makes the practice more interesting.”* P3 recounted:

The feature I found most interesting was the part where the AI reads the questions. Especially in Speech AI, it displays an image, then reads the question and gives me time to prepare, just like a real interview. That’s why I found it fascinating. (P3)

Some students felt a strong sense of satisfaction when they noticed improvement. P5 mentioned *“one situation where the AI tool made me feel more confident was during my most recent practice session, where there were fewer red marks on the screen. I felt that I had improved and was very satisfied with my progress,”* and P2 reported, *“even though I was just speaking based on my own thoughts, my pronunciation naturally improved. By the final exercises, my pronunciation score was quite good. That made me feel happy and pleased.”*

On the other hand, a few negative emotional reactions were observed. Firstly, P1 admitted feeling nervous when responding to unfamiliar questions.

There were times when the AI asked questions that I hadn't thought about before, and in those moments, I felt a bit nervous and took time to think. When I started speaking, I felt anxious and worried, so I ended up unable to respond. (P1)

Secondly, P4 and P5 questioned the reliability of the feedback, as P4 explained that *"I had many concerns and worries because I thought AI was just a programmed system, and I was not sure if it could address all aspects of speaking and the mistakes I had in speaking"*. P5 further remarked:

AI's feedback follows a fixed pattern, and it always evaluates in the same way. Also, when it comes to voice quality, sometimes the feedback feels a bit one-sided. This can make me feel uncertain, I do not know if I'm actually wrong or if I made a mistake and I was correct. (P5)

Finally, in her journal, P3 shared that her confidence sometimes declined due to receiving low scores despite her strong efforts to "read that part many times". She mentioned feeling *"sometimes I felt discouraged and lost confidence when I thought I'd tried hard to pronounce a part correctly, but the score still turned out quite low"*. However, P3 recognized that this might have been because the AI-powered platforms were designed to identify "many pronunciation errors."

Behavioral Component – Learning Habits and Engagement

Behaviorally, students demonstrated high levels of engagement and a willingness to continue using AI tools. Some described their efforts to repeatedly practice errors words or sounds. For instance, P2 shared, *"When I received feedback from the AI, I usually practiced only the word that I got wrong. I repeated those mispronounced words several times before answering the question again."* In this vein, P3 shared that "I occasionally skip certain words which I believe I have pronounced correctly, instead of repeating the practice ones which the website highlights as mistakes."

Additionally, participants also noted changes in their learning habits. P5 admitted, *"I have practiced pronunciation more frequently. In the past, I only looked up how to pronounce a word when I needed to say it correctly. But now, since I want to communicate with the AI, I have to practice pronunciation more often, which motivates me more than before."* Furthermore, all five participants expressed a desire to continue using AI tools beyond the study period. As P1 and P4 expressed:

Since practicing helps me improve, I can also enhance my scores at school. So, I will continue using it to speak better. (P1)

When using these tools, I find that I can gain many benefits. However, there are still some challenges. I hope that as AI improves and addresses its limitations, I will continue using these tools as part of my daily practice in the future. (P4)

Discussion

Students' Perceptions of Benefits and Challenges

The results demonstrate that students generally view AI tools as beneficial, particularly in improving segmental pronunciation. Quantitative data revealed high agreement on the usefulness of AI in practicing vowel and consonant sounds. Long vowels received the highest rating, followed by short vowels and other consonants. These results echo findings from previous studies (Anggraini, 2022; Dennis, 2024; Marlinda & Huda, 2024; Nguyen et al., 2025; Nguyen et al., 2025), reporting improved pronunciation following AI-assisted training. Similarly, studies by Hoang et al. (2023), Amin (2024), and

Nguyen and Nguyen (2024) confirm that AI tools help students better distinguish segmental features like vowels and consonants.

In contrast, the perceived impact of AI tools on suprasegmental features was more modest. Word stress and rhythm were rated higher than intonation, connected speech, and chunking/pauses. This limited support may stem from the complexity of suprasegmental features, which require subtle and contextual feedback that current AI systems are not fully equipped to deliver. While some studies (e.g., Yoo & Ahn, 2024) acknowledged that AI could enhance rhythm and stress, students in this study felt that it did not provide sufficient support for connected speech or intonation. This concern was also noted by Senowarsito and Ardini (2023), who emphasized AI's limitations in evaluating prosodic features.

Despite the benefits, students encountered several challenges. The most frequently reported issue was unclear feedback. Participants described how AI tools often highlighted entire sentences instead of pinpointing specific errors, making it difficult to identify and address pronunciation issues effectively. This finding is consistent with Mohammadkarimi's (2024) exploration and suggestion that many students struggle to understand AI-generated feedback, so they need a better user interface design.

Technical limitations were another key concern. Students reported inaccuracies in speech recognition, such as correct pronunciation being marked incorrect. These inaccuracies disrupted the learning experience and undermined learners' confidence in the tools, corroborating observations by Hoang et al. (2023), Mohammadkarimi (2024), Nguyen et al. (2025), and Nguyen et al. (2025). Additionally, limited feedback on suprasegmental aspects left learners unsure of how to improve in areas like intonation and connected speech.

In sum, while previous research has acknowledged general benefits of AI in pronunciation training, this study provides more nuanced insights. It confirms AI's strength in segmental features while highlighting persistent gaps in suprasegmental feedback. This dual focus offers a valuable contribution to the field by clarifying both the potentials and limitations of current AI-based pronunciation support.

Students' Attitudes Towards AI Tools

Cognitive Attitudes

Cognitively, students viewed AI tools as effective in identifying errors, facilitating self-correction, and tracking gradual improvement, especially in segmental pronunciation. This aligns with research by Sun (2023), Dennis (2024), Mohammadkarimi (2024), and Nguyen et al. (2025), who found that learners value the immediate, targeted feedback provided by AI systems. Unlike earlier studies that focused broadly on effectiveness, this study also revealed that students actively monitored their progress over time. They noted improvements in specific areas such as final sounds and subtle consonantal contrasts, offering a more detailed account of learner awareness than typically found in the literature. According to Rosenberg and Hovland (1960), the cognitive component of attitude involves beliefs, thoughts, and knowledge about an object, which explains why learners' perception of AI as accurate and helpful played a central role in shaping their overall attitude toward its use.

Importantly, students also reported that AI tools motivated them to practice more frequently. The perceived benefits, such as real-time feedback and score tracking, prompted sustained engagement. This supports findings from Hoang et al. (2023), Sun (2023), Abimanto and Sumarsono (2024), and Nguyen et al. (2025), who argued that AI can foster self-directed learning by encouraging learners to take ownership of their development. This is consistent with the Tripartite Model's view that cognitive appraisals, such as recognizing the usefulness of a tool, can influence motivation and future behavioral

intentions (Rosenberg & Hovland, 1960). It should be noted that the current study is an extension of the Tripartite Model as it revealed evidence of feature specificity in artificial intelligence's contribution to students' pronunciation learning by demonstrating that perceived usefulness contributed to an increase in students' behavioral intention to practice their pronunciation, although this support did not occur uniformly for all domains involving students' pronunciation since limited support through artificial intelligence was provided for suprasegmental speech features such as connected speech and chunking. By illuminating how AI promotes not only correction but also sustained motivation, the current study underscores the cognitive depth of student engagement.

Affective Attitudes

Emotionally, the study revealed a mixture of positive and negative reactions. Many students expressed enjoyment and satisfaction with AI-based practice, citing the variety of practice topics, interactive features, and noticeable improvement in performance. These findings are consistent with those of Anggraini (2022), Vu (2024), and Nguyen et al. (2024), who have emphasized the motivational power of engaging digital tools. The participants noticed their improvement, which was explained by Mohammadkarimi (2024) that students felt more self-assured after using AI for pronunciation training. Rosenberg and Hovland (1960) defined the affective component as involving emotional responses toward an object, which helps explain how feelings of enjoyment, pride, or anxiety influence students' attitudes toward AI tools.

However, this study uncovered an emotional nuance not widely reported in prior research: initial anxiety. Some participants felt nervous when first using AI tools, especially when facing unexpected or challenging prompts. These emotional responses suggest that while AI can support confidence-building over time, initial orientation and support are critical to help students overcome hesitation. Another emotional barrier involved uncertainty about the accuracy and fairness of AI evaluations. Some students expressed doubt about the validity of the feedback, especially when receiving contradictory or unclear corrections. This concern reflects previous findings by Mohammadkarimi (2024), who observed that rigid feedback mechanisms can make students' question the objectivity of AI evaluations. Another finding involved a particular case where the participant noticed a decline in confidence after receiving low scores despite repeated practice, which reflected the tools' built-in capability to detect pronunciation errors.

Behavioral Attitudes

Behaviorally, students demonstrated consistent engagement with AI tools. They frequently practiced and revisited incorrect responses, showing a proactive approach to refining their pronunciation. This aligns with the work of Abimanto and Sumarsono (2024), who found that students benefited from repeated, focused practice with AI support. The behavioral component, as proposed by Rosenberg and Hovland (1960), refers to the actions or observable responses toward an object; thus, students' increased usage, repeated practice, and intention to continue using AI illustrate a clear behavioral manifestation of their overall positive attitude.

Notably, students described a shift in their learning habits. One student reflected that she used to check pronunciation only when needed, but had adopted more regular and deliberate practice after using AI. Most participants expressed a clear intention to continue using AI tools in the future, recognizing their value in providing personalized feedback and supporting long-term learning. These findings are consistent with Amin (2024), who reported students' willingness to integrate text-to-speech tools into their ongoing language development.

However, this study adds a critical contribution by clarifying the reasoning behind this behavioral shift. It was not merely convenience or novelty that prompted continued use, but rather students' recognition of tangible improvements in pronunciation. This aligns with the Tripartite Model's assertion that consistent behaviors often emerge when both cognitive evaluations (e.g., perceived usefulness) and affective responses (e.g., enjoyment or confidence) reinforce each other (Rosenberg & Hovland, 1960). This insight emphasizes the importance of perceived effectiveness in fostering sustainable behavioral change.

Conclusion

This study explored English majors' perceptions and attitudes toward using AI-powered tools to enhance their pronunciation, focusing on both segmental and suprasegmental features. The findings revealed that students generally found AI tools effective in supporting segmental pronunciation, especially with long vowels. However, perceived improvements in suprasegmental features, such as pauses, speech chunking, and connected speech, were limited. Challenges reported included unclear feedback and technical inaccuracies, particularly in recognizing suprasegmental elements. Despite these issues, participants expressed positive cognitive evaluations, emotional engagement, and behavioral commitment toward continued use of AI tools. The Tripartite Model of Attitudes (Rosenberg & Hovland, 1960) was instrumental in illustrating how students' beliefs, feelings, and behaviors shaped their experience.

Theoretically, this study contributes to the field by applying an attitudinal framework to examine students' engagement with AI in pronunciation training, an area that remains underexplored in the literature. The use of a mixed-methods design incorporating questionnaires, semi-structured interviews, and journal diaries enabled a nuanced understanding of students' experiences and self-perceptions. This approach enriches research on technology-assisted pronunciation learning by distinguishing the effects on segmental and suprasegmental features. Practically, the findings highlight both the potential and limitations of current AI tools, emphasizing their effectiveness for segmental training while identifying gaps in suprasegmental support.

The pedagogical importance of these results lies in the implication that teachers use AI systems in different ways to develop students' segmental and suprasegmental features. For segmental aspects, teachers are recommended to use AI systems to help students with those sounds where there was a subjective perception of greatest improvement, such as long vowels and challenging contrastive sounds such as /ɪ/ vs. /i:/ and θ/ vs. s/. However, teachers are also recommended to encourage students to use AI systems in a critical way to at least verify the occurrence and patterns of error with speech sounds with activities such as comparing recordings with those of teachers or native speakers and checking up incorrectly recognized speech with teacher explanations. For suprasegmental aspects, AI systems were seen as more supportive for developing stress and rhythmic systems, although further teacher involvement and tasks are required to guide students to use AI systems to present and work with speech in such areas as pausing, chunking, and running conversation.

Several areas for further research emerge from this study's limitations. To enhance qualitative reliability, future studies should incorporate multiple raters and interrater agreement checks. Expanding data collection methods to include observations or controlled experiments would offer richer insights into actual learner interactions with AI. Additionally, employing randomized and diverse sampling across institutions would improve generalizability. Finally, future validation of measurement instruments through psychometric statistics is necessary to strengthen the reliability and validity of the questionnaire.

References

- Abimanto, D., & Sumarsono, W. (2024). Improving English pronunciation with AI speech-recognition technology. *Acitya: Journal of Teaching and Education*, 6(1), 146–156. <https://doi.org/10.30650/ajte.v6i1.3810>
- Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, 52(1), 27–58. <https://doi.org/10.1146/annurev.psych.52.1.27>
- Albiladi, W. S. (2019). Teaching English pronunciation revisited: The challenges of teaching EFL in non-English-speaking countries. *European Journal of Foreign Language Teaching*, 4(3), 41–50. <https://doi.org/10.5281/zenodo.3566658>
- Amin, E. A.-R. (2024). EFL students' perception of using AI text-to-speech apps in learning pronunciation. *Migration Letters*, 21(3), 887–903. <https://migrationletters.com/index.php/ml/article/view/6905>
- Anggraini. (2022). Improving students' pronunciation skills using ELSA Speak Application. *Journal of English Language and Pedagogy*, 5(1), 135–141.
- Babaeian, A. (2023). Pronunciation assessment: Traditional vs modern modes. *Journal of Education for Sustainable Innovation*, 1(1), 61–68. <https://doi.org/10.56916/jesi.v1i1.530>
- Bakar, Z. A., & Abdullah, M. R. T. L. (2015). Importance of correct pronunciation in spoken English: Dimension of second language learners' perspective. *Pertanika Journal of Social Sciences and Humanities*, 23(S), 143–158. <http://www.pertanika.upm.edu.my/pjssh/browse/special-issue?article=JSSH-S0023-2015>
- Busso, A., & Sanchez, B. (2024). Advancing communicative competence in the digital age: A case for AI tools in Japanese EFL education. *Technology in Language Teaching & Learning*, 6(3), 1211. <https://doi.org/10.29140/tltl.v6n3.1211>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, A. (2014). *Pronunciation and phonetics: A practical guide for English language teachers*. Routledge.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide* (2th ed.). Cambridge University Press.
- Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dennis, N. K. (2024). Using AI-powered speech recognition technology to improve English pronunciation and speaking skills. *IAFOR Journal of Education: Technology in Education*, 12(2), 107. <https://doi.org/10.22492/ije.12.2.05>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company. <https://doi.org/10.1075/llt.42>
- Dhanapal, C., Asharudeen, N., & Alfaruque, S. Y. (2024). Impact of artificial intelligence versus traditional instruction for language learning: A survey. *World Journal of English Language*, 14(2), 182–193. <https://doi.org/10.5430/wjel.v14n2p182>
- Fraenkel, J., Wallen, N., & Hyun, H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill Education.
- Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, 2(3), 119–128. https://ijhss.thebrpi.org/journals/Vol_2_No_3_February_2012/17.pdf

- Gilakjani, A. P. (2016). English pronunciation instruction: A literature review. *International Journal of Research in English Education*, 1(1), 1–6.
- Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Pearson.
- Hoang, N. T., Duong, N. H., & Le, D. H. (2023). Exploring Chatbot AI in improving vocational students' English pronunciation. *AsiaCALL Online Journal*, 14(2), 140–155. <https://doi.org/10.54855/acoj.231429>
- Jenkins, J. (2000). *The phonology of English as an international language* (Oxford Applied Linguistics). Oxford University Press.
- Junaidi, Hamuddin, B., Julita, K., Rahman, F., & Derin, T. (2020). Artificial Intelligence in EFL context: Rising students' speaking performance with Lyra Virtual Assistance. *International Journal of Advanced Science and Technology*, 29(5), 6735–6741. <http://repository.unhas.ac.id:443/id/eprint/11102>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Khan, M. O., Nazim, M., & Alzubi, A. A. F. (2024). Exploring Arab EFL learners' attitudes: Is Grammarly a game-changer in academic writing classes? *Educational Administration: Theory and Practice*, 30(4), 1365–1371. <https://doi.org/10.53555/kuey.v30i4.1612>
- Kobilova, N. R. (2022). Importance of pronunciation in English language communication. *Academic Research in Educational Sciences*, 3(6). <https://ares.uz/storage/app/uploads/public/62b1db16b62b1db16bcc76435384219.pdf>
- Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed.). Cengage Learning.
- Lee, B. J. (2025). Reimagining language learning: AI-driven innovations for engagement and growth. *Technology in Language Teaching & Learning*, 6(3), 102773. <https://doi.org/10.29140/tltl.v6n3.102773>
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377. <https://doi.org/10.2307/3588485>
- Marlinda, S., & Huda, N. N. (2024). Improving pronunciation skills with ELSA Speak an AI-based learning. *Jiic: Jurnal Intelek Insan Cendikia*, 1(6), 1866–1871. <https://jicnusantara.com/index.php/jiic/article/view/765/861>
- 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. <https://doi.org/10.24071/llt.v27i1.8151>
- Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. <https://doi.org/10.1017/S0958344023000113>
- Nguyen, X. H., Le, T. T., Do, Đ. K., Hoang, Q. N., & Nguyen, T. T. H. (2024). Students' attitudes toward utilizing AI-based technologies to improve their speaking proficiency. *Ho Chi Minh City University of Education Journal of Science*, 21(5), 957. [https://doi.org/10.54607/hcmue.js.21.5.4270\(2024\)](https://doi.org/10.54607/hcmue.js.21.5.4270(2024))
- Nguyen, K. N., & Nguyen, M. T. T. (2024). Tertiary students' perceptions on ELSA Speak application for pronunciation learning. *European Journal of Applied Linguistics Studies*, 7(1). <https://doi.org/10.46827/ejals.v7i1.483>
- Nguyen, N. V., Vo, T. T., & Tran, V. D. T. (2025). AI-driven pronunciation assessment: The impact of SpeechAce on EFL learners' pronunciation competency. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ)*, 26(3), 84–106. <https://doi.org/10.54855/callej.252635>
- Nguyen, T. S., Nguyen, T. D. T., Hoang, N. Q. N., & Do, T. K. H. (2025). How AI-powered voice recognition has supported pronunciation competence among EFL university learners. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ)*, 26(3), 64–83. <https://doi.org/10.54855/callej.252634>

- Otlowski, M. (1998). Pronunciation: What are the expectations? *The Internet TESL Journal*, 4(1), 115–129. <http://iteslj.org/Articles/Otlowski-Pronunciation.html>
- Phi, T. T. T., & Mai, L. (2024). English-majored students' attitudes towards English pronunciation and challenges of acquiring proper pronunciation: A case study of Thang Long University. *Tạp chí Khoa học Thăng Long: Khoa học Ứng dụng*, 3(2), 49–67. <https://science.thanglong.edu.vn/index.php/vola/article/view/140>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective and behavioral components of attitudes. In M. J. Rosenberg & C. I. Hovland (Eds.), *Attitude organization and change: An analysis of consistency among attitude components*. Yale University Press.
- Sánchez, L., Morales, A., & Rodríguez, I. (2024). Speech recognition software as a tool to enhance EFL learners' pronunciation. *Lengua y Sociedad*, 23(2), 964–983. <https://doi.org/10.15381/lengsoc.v23i2.26074>
- Selladurai, S., & Pandey, P. S. (2024). The role of AI tools in polishing English pronunciation skills for non-native speakers. *International Journal of Communication Networks and Information Security (IJCNIS)*, 16(1), 1455–1467. <https://www.ijcnis.org/index.php/ijcnis/article/view/7190>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tran, H. C. (2024). Using AI for students' attitudes to improve their speaking skills at a university. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(4), 557–562. <https://doi.org/10.62225/2583049X.2024.4.4.3070>
- Vančová, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education*, 11(3), 12–24. <https://intapi.sciendo.com/pdf/10.2478/jolace-2023-0022>
- Vu, L. T. (2024). Nghiên cứu áp dụng phần mềm ELSA để cải thiện phát âm cho sinh viên Trường Đại học Giáo dục, Đại học Quốc gia Hà Nội. *Journal of Educational Equipment: Applied research*, 2(311), 317–318. <https://scholar.dlu.edu.vn/thuvienso/bitstream/DLU123456789/290235/1/97356-925-205991-1-2-20240607.pdf>
- Yaghma, F. (2003). Content validity and its estimation. *Journal of Medical Education*, 3(1), 25–27. <https://doi.org/10.22037/jme.v3i1.870>
- Yates, L., & Zielinski, B. (2009). *Give it a go: Teaching pronunciation to adults*. Macquarie University.
- Yoo, S., & Ahn, H. (2024). The effects of prosody training with AI Chatbot on the English pronunciation improvement of Korean EFL learners. *Korean Journal of English Language and Linguistics*, 24, 1300–1317. <https://doi.org/10.15738/kjell.24..202411.1300>
- Zhao, Y. (2024). The exploration of artificial intelligence in pronunciation teaching. In *Proceedings of the International Conference on Data Science and Engineering* (pp. 127–131). <https://doi.org/10.5220/0012829100004547>
- Zuhairya, N., Maharani, P., Ridwan, A., Khairunnisa, Ahwani, S., & Lubis, Y. (2024). The role of suprasegmental features in English phonology: Prosodic hierarchy and intonation patterns. *Fonologi: Jurnal Ilmuan Bahasa Dan Sastra Inggris*, 2(2), 154–161. <https://doi.org/10.61132/fonologi.v2i2.670>

Appendices

Appendix A. Questionnaire

The purpose of this survey is to gather data to support the completion of my thesis. Your participation is highly valued, and I kindly request you to take a few moments to complete the survey. **I assure that all personal information collected will remain strictly confidential and will be used solely for academic research purposes.** Your insights and feedback are greatly appreciated and will contribute significantly to the success of this research.

Thank you for your time and cooperation.

For better understanding the theoretical information in the survey, please carefully read the following sentences.

- Pronunciation includes two main features: segmental and suprasegmental (Celce-Murcia et al., 2010). The segmental features include individual sounds, such as vowels and consonants; suprasegmental features include features beyond the sounds; such as pitch, intonation, stress, and rhythm (Ladefoged & Johnson, 2011).
- Nowadays, there are many AI tools (ELSA Speak, Lingoland, SpeechAce, etc.) applied to language learning, especially pronunciation practice. This survey aims to explore the opinions of language learners towards using AI to improve pronunciation.

Part 1: Demographics

1. What is your gender? (*Giới tính của bạn là gì?*)
 - Male (*Nam*)
 - Female (*Nữ*)
2. How would you rate your current level of English pronunciation? (*Bạn đánh giá trình độ phát âm tiếng Anh hiện tại của mình như thế nào?*)
 - Beginner (A1) (*Căn bản*)
 - Elementary (A2) (*Sơ cấp*)
 - Intermediate (B1) (*Trung cấp*)
 - Upper-Intermediate (B2) (*Trung cao cấp*)
 - Advanced (C1) (*Cao cấp*)
 - Proficient (C2) (*Thành thạo*)
3. How often do you practice English pronunciation? (*Bạn có thường xuyên luyện phát âm tiếng Anh không?*)
 - Daily (at least 1 session every day) (*Hàng ngày (ít nhất 1 buổi mỗi ngày)*)
 - Several times a week (3-6 days per week) (*Vài lần một tuần (3-6 ngày/tuần)*)
 - Once or twice a week (1-2 days per week) (*1 hoặc 2 lần một tuần (1-2 ngày/tuần)*)
 - Occasionally (less than once a week) (*Thỉnh thoảng (ít hơn 1 lần 1 tuần)*)
 - Rarely or never (*Hiếm khi hoặc không bao giờ*)

Part 2: Experience with AI Tools in Pronunciation

1. Please name some AI tools to practice pronunciation you have used (e.g. ELSA Speak, Lingoland, SpeechAce, etc.). (Open-ended) (*Hãy kể tên một số công cụ AI luyện phát âm mà bạn đã sử dụng (Câu hỏi mở)*)
-

2. How often do you use AI tools for pronunciation practice? (*Bạn có thường xuyên sử dụng công cụ AI để luyện phát âm không?*)

- Daily (at least 1 session every day) (*Hàng ngày (ít nhất 1 buổi mỗi ngày)*)
- Several times a week (3-6 days per week) (*Vài lần một tuần (3-6 ngày/tuần)*)
- Once or twice a week (1-2 days per week) (*1 hoặc 2 lần một tuần (1-2 ngày/tuần)*)
- Occasionally (less than once a week) (*Thỉnh thoảng (ít hơn 1 lần 1 tuần)*)
- Rarely or never (*Hiếm khi hoặc không bao giờ*)

Please rate your opinion with the following statements with the provided scale (*Vui lòng đánh giá ý kiến của bạn với các nhận định sau theo thang điểm được cung cấp*):

- (1) Strongly disagree (*Hoàn toàn không đồng ý*)
- (2) Disagree (*Không đồng ý*)
- (3) Neutral (*Trung lập*)
- (4) Agree (*Đồng ý*)
- (5) Strongly agree (*Hoàn toàn đồng ý*)

Theme 1: Segmental features

Chủ đề 1: Đặc điểm âm vị đoạn tính

What are your opinions towards using AI to improve your segmental features of pronunciation? (*Ý kiến của bạn về việc sử dụng AI để cải thiện các đặc điểm âm vị đoạn tính trong phát âm của bạn là gì?*)

	(1)	(2)	(3)	(4)	(5)
1. AI tools can help me improve my pronunciation with short vowel sounds (/ɪ/, /æ/, /e/, /a/, /ʌ/, /ʊ/, /ə/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm ngắn (/ɪ/, /æ/, /e/, /a/, /ʌ/, /ʊ/, /ə/).</i>					
2. AI tools can help me improve my pronunciation with long vowel sounds (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm dài (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/).</i>					
3. AI tools can help me improve my pronunciation with diphthongs (/ɪə/, /eə/, /eɪ/, /ɔɪ/, /aɪ/, /əʊ/, /aʊ/, /ʊə/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm đôi (/ɪə/, /eə/, /eɪ/, /ɔɪ/, /aɪ/, /əʊ/, /aʊ/, /ʊə/).</i>					
4. AI tools can help me improve my pronunciation with triphthongs (/aʊə/, /aɪə/, /eɪə/, /əʊə/, /ɔɪə/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm ba (/aʊə/, /aɪə/, /eɪə/, /əʊə/, /ɔɪə/).</i>					
5. AI tools can help me improve my pronunciation with voiced consonant sounds (/b/, /d/, /g/, /v/, /z/, /ð/, /dʒ/, /ʒ/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm hữu thanh (/b/, /d/, /g/, /v/, /z/, /ð/, /dʒ/, /ʒ/).</i>					
6. AI tools can help me improve my pronunciation with voiceless consonant sounds (/p/, /t/, /k/, /f/, /s/, /θ/, /tʃ/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm vô thanh (/p/, /t/, /k/, /f/, /s/, /θ/, /tʃ/).</i>					
7. AI tools can help me improve my pronunciation with other consonant sounds (/m/, /n/, /ŋ/, /l/, /j/, /h/, /r/, /w/). <i>Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm khác (/m/, /n/, /ŋ/, /l/, /j/, /h/, /r/, /w/).</i>					

8. Why do you think AI tools are (or aren't) effective in improving segmental features? (Open-ended) *(Tại sao bạn nghĩ các công cụ AI có (hoặc không) hiệu quả trong việc cải thiện các đặc điểm âm vị đoạn tính?) (Câu hỏi mở)*
-
9. What are the challenges you face when using AI tools to practice segmental features? (You can choose more than 1) *(Những thách thức bạn gặp phải khi sử dụng các công cụ AI để luyện tập các đặc điểm âm vị đoạn tính là gì?) (Có thể chọn nhiều hơn 1)*
- The AI tool mis-recognized some of my vowel/consonant sounds.
 - Lack of clear feedback on specific sound errors.
 - Difficulty in adjusting sounds even after receiving feedback.
 - The AI doesn't cover all types of sounds (e.g., regional accents).
 - Technical issues (e.g., software bugs, slow Internet)
 - None of the above
10. Are there any other challenges you face when using AI tools to practice segmental features? If yes, please clarify them. (Open-ended) *(Bạn có gặp phải thách thức nào khác khi sử dụng các công cụ AI để luyện tập các đặc điểm âm vị đoạn tính không? Nếu có, xin vui lòng làm rõ chúng.) (Câu hỏi mở)*
-

Theme 2: Suprasegmental features

Chủ đề 2: Đặc điểm siêu âm vị đoạn tính

What are your opinions towards using AI to improve your suprasegmental features of pronunciation? *(Ý kiến của bạn về việc sử dụng AI để cải thiện các đặc điểm siêu âm vị đoạn tính trong phát âm của bạn là gì?)*

	(1)	(2)	(3)	(4)	(5)
1. AI tools can help me improve my pronunciation of individual words with the correct stress. <i>Các công cụ AI có thể giúp tôi cải thiện cách phát âm từng từ với trọng âm đúng.</i>					
2. AI tools can help me improve my pronunciation of complete sentences with correct intonation. <i>Các công cụ AI có thể giúp tôi cải thiện cách phát âm câu hoàn chỉnh với ngữ điệu đúng.</i>					
3. AI tools can help me improve my pronunciation in connected speech (elision, linking, assimilation, intrusion, etc.). <i>Các công cụ AI có thể giúp tôi cải thiện cách phát âm trong lời nói liên kết (lược âm, nối âm, đồng hóa, chèn âm,...)</i>					
4. AI tools can help me improve my use of pauses and speech chunking. <i>Các công cụ AI có thể giúp tôi cải thiện việc sử dụng ngắt câu và phân đoạn trong lời nói.</i>					
5. AI tools can help me improve my rhythm in speaking English. <i>Các công cụ AI có thể giúp tôi cải thiện nhịp điệu khi nói tiếng Anh.</i>					

6. Why do you think AI tools are (or aren't) effective in improving suprasegmental features? (Open-ended) *(Tại sao bạn nghĩ các công cụ AI có (hoặc không) hiệu quả trong việc cải thiện các đặc điểm siêu âm vị đoạn tính?) (Câu hỏi mở)*
-
7. What are the challenges you face when using AI tools to practice suprasegmental features? (You can choose more than 1) *(Những thách thức bạn gặp phải khi sử dụng các công cụ AI để luyện tập các đặc điểm siêu âm vị đoạn tính là gì?) (Có thể chọn nhiều hơn 1)*
- Limited features for improving intonation or rhythm *(Các tính năng hạn chế để cải thiện ngữ điệu hoặc nhịp điệu)*
 - The AI tool doesn't accurately recognize stress/intonation issues. *(Công cụ AI không nhận dạng chính xác các vấn đề về trọng âm/ngữ điệu.)*
 - I find it difficult to improve my intonation based on the feedback provided. *(Tôi cảm thấy khó cải thiện ngữ điệu của mình dựa trên những phản hồi được cung cấp.)*
 - AI tools don't provide real-time feedback on rhythm or pauses. *(Các công cụ AI không cung cấp phản hồi theo thời gian thực về nhịp điệu hoặc khoảng dừng.)*
 - The AI can't fully understand sentence stress in spontaneous speech. *(AI không thể hiểu đầy đủ trọng âm của câu trong lời nói tự phát.)*
 - Technical issues (e.g., software bugs, slow Internet) *(Sự cố kỹ thuật (ví dụ: lỗi phần mềm, Internet chậm))*
 - None of the above *(Không có điều nào ở trên)*
8. Are there any other challenges you face when using AI tools to practice segmental features? If yes, please clarify them (Open-ended) *(Bạn có gặp phải thách thức nào khác khi sử dụng các công cụ AI để luyện tập các đặc điểm siêu âm vị đoạn tính không? Nếu có, xin vui lòng làm rõ chúng.) (Câu hỏi mở)*
-
9. How often do you encounter technical issues (e.g., software bugs, slow Internet) when using AI tools for pronunciation practice? *(Bạn có thường xuyên gặp phải sự cố kỹ thuật (ví dụ: lỗi phần mềm, Internet chậm) khi sử dụng các công cụ AI để luyện phát âm không?)*
- Very often *(Rất thường xuyên)*
 - Often *(Thường xuyên)*
 - Sometimes *(Thỉnh thoảng)*
 - Rarely *(Hiếm khi)*
 - Never *(Không bao giờ)*
10. Do you think these challenges affect your ability to improve your overall pronunciation? Why? (Open-ended) *(Bạn có nghĩ những thử thách này ảnh hưởng đến khả năng cải thiện khả năng phát âm tổng thể của bạn không? Tại sao?) (Câu hỏi mở)*
-

Appendix B. Source and Development of the Questionnaire

The questionnaire used in this study was self-designed because no existing tool adequately captured the constructs required to answer our research questions in the given cultural and contextual setting. Development followed a systematic, multi-stage process to ensure both content validity and response clarity. The steps included (1) item generation based on theory and research question, (2) expert review, (3) cognitive interviews with target participants, and (4) preliminary psychometric testing. Details of each stage are summarized below, with supporting evidence provided in the accompanying tables and figures.

1. Self-designed instrument

The questionnaire was developed specifically for this study to address Research Question 1. Item generation was guided by theories of segmental and suprasegmental (Celce-Murcia et al., 2010), ensuring coverage of the relevant constructs. Existing tools were reviewed but judged unsuitable for the study context.

2. Expert review

A panel of four experts in TESOL reviewed the draft items for clarity, relevance, and representativeness. We calculated item-level CVI and S-CVI/Ave; items below threshold were revised accordingly.

Table 1 *Expert Demographics*

Number	Title – Highest degree	Gender	Affiliation based in
1	Associate Professor- PhD	Male	Thailand
2	Associate Professor- PhD	Male	Vietnam
3	PhD	Female	Vietnam
4	PhD	Male	Thailand
5	PhD	Female	U.S.A
6	PhD	Female	Vietnam

Table 2 *Expert Ratings and I-CVI Calculation*

Item	Expert 1	Expert 2	Expert 3	Expert 4	Expert in Agreement	I-CVI
Q1	1	1	1	1	4	1
Q2	1	1	1	1	4	1
Q3	1	1	1	1	4	1
Q4	1	1	1	1	4	1
Q5	1	1	1	1	4	1
Q6	1	1	1	1	4	1
Q7	1	1	1	1	4	1
Q8	1	1	1	1	4	1
Q9	1	1	1	1	4	1
Q10	1	1	1	1	4	1
Q11	1	1	1	1	4	1
Q12	1	1	1	1	4	1



Khánh Nguyễn Hữu Lam <2157010297khanh@ou.edu.vn>

to

Oct 24, 2024, 6:49 PM ☆ ↶ ⓘ

Dear Dr

I hope this message finds you well. My name is Nguyễn Hữu Lam Khánh, and I am conducting a research project under supervision by Mr. Tran Thanh Tan, which focuses on senior English majors' perceptions toward benefits and challenges of using AI tools to improve English pronunciation.

Our scale is built upon (1) theory of English pronunciation with two features, namely segmental and suprasegmental (Celce-Murcia et al., 2010); (2) the tripartite model of attitudes (Rosenberg & Hovland, 1960). Specifically, we have created a questionnaire-item instrument, grouped into 2 themes:

- Segmental features
- Suprasegmental features

To ensure the content validity of our instrument, we are seeking feedback from distinguished experts like yourself in the field of English pronunciation. As introduced by Mr. Tan Tran, we believe that your valuable insights and expertise would greatly contribute to refining the scale. Specifically, we would appreciate your feedback on the clarity, relevance, and completeness of the items, as well as any suggestions you may have for improvement.

Your participation in this review process would be instrumental in helping us ensure that the scale is both theoretically sound and practically applicable in diverse educational settings. We hope you would be willing to lend your expertise to this endeavor.

The pool of items is included in the link:

https://docs.google.com/document/d/1wmyU0Cyl-HO_OsXeoB1130nTtaGXkDK_2eK2rzTwp2/edit?usp=sharing

Your feedback would be greatly appreciated, and we are happy to accommodate any timeline that suits your availability.

Thank you for considering this request. I look forward to the possibility of collaborating with you on this important research initiative.

Warm regards,

Figure 1 Screenshot of the Exchange Email with One Expert

4. Cognitive interviews

Five target participants were recruited to assess comprehension using think-aloud and probing techniques. Feedback was documented in a structured log and informed revisions to wording, examples, and item order.

Table 4 Participant Demographics

Number	Year of Study	Sex
1	Second year	Female
2	Second year	Female
3	Second year	Female
4	Second year	Female
5	Second year	Female

	A	B	C	D	E	F	G	H	I	J	K
			English items	Vietnamese item	S1	S2	S3	S4	S5	Revised items (English)	Revised items (Vietnamese)
1			1. AI tools can help me improve my pronunciation with short vowel sounds (/ɪ/, /æ/, /e/, /ɑ/, /ʊ/, /u/)	1. Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm ngắn (/ɪ/, /æ/, /e/, /ɑ/, /ʊ/, /u/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
2			2. AI tools can help me improve my pronunciation with long vowel sounds (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/)	2. Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm dài (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
3			3. AI tools can help me improve my pronunciation with diphthongs (/aɪ/, /eɪ/, /ɪə/, /aʊ/, /əʊ/, /aʊ/, /əʊ/)	3. Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm đôi (/aɪ/, /eɪ/, /ɪə/, /aʊ/, /əʊ/, /aʊ/, /əʊ/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
4			4. AI tools can help me improve my pronunciation with triphthongs (/aʊə/, /aɪə/, /eɪə/, /əʊə/, /əʊə/)	4. Các công cụ AI có thể giúp tôi cải thiện phát âm các nguyên âm ba (/aʊə/, /aɪə/, /eɪə/, /əʊə/, /əʊə/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
5			5. AI tools can help me improve my pronunciation with voiced consonant sounds (/b/, /d/, /g/, /v/, /z/, /ð/, /dʒ/, /ʒ/)	5. Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm hữu thanh (/b/, /d/, /g/, /v/, /z/, /ð/, /dʒ/, /ʒ/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
6			6. AI tools can help me improve my pronunciation with voiceless consonant sounds (/p/, /t/, /k/, /f/, /s/, /θ/, /tʃ/, /ʃ/)	6. Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm vô thanh (/p/, /t/, /k/, /f/, /s/, /θ/, /tʃ/, /ʃ/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
7			7. AI tools can help me improve my pronunciation with other consonant sounds (/m/, /n/, /ŋ/, /l/, /j/, /h/, /r/, /w/)	7. Các công cụ AI có thể giúp tôi cải thiện phát âm các phụ âm khác (/m/, /n/, /ŋ/, /l/, /j/, /h/, /r/, /w/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
8			1. I find it difficult to use AI tools to improve my pronunciation with short vowel sounds (/ɪ/, /æ/, /e/, /ɑ/, /ʊ/, /u/)	1. Tôi cảm thấy khó khăn khi sử dụng các công cụ AI để cải thiện phát âm các nguyên âm ngắn (/ɪ/, /æ/, /e/, /ɑ/, /ʊ/, /u/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
9			2. I find it difficult to use AI tools to improve my pronunciation with long vowel sounds (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/)	2. Tôi cảm thấy khó khăn khi sử dụng các công cụ AI để cải thiện phát âm các nguyên âm dài (/i:/, /u:/, /ɔ:/, /ɜ:/, /ɑ:/)	no comment	no comment	no comment	no comment	no comment	No revision	No revision
10											

Figure 2 Cognitive Interview Log (Excel screenshot)

4. Preliminary psychometric check

A pilot dataset was analyzed. The KMO statistic was 0.934 and Bartlett's test was significant ($p < .001$), supporting factorability for further analysis.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.934
Bartlett's Test of Sphericity	Approx. Chi-Square	1164.347
	df	66
	Sig.	.000

Figure 3 SPSS Output Screenshot

5. Reliability analysis (Cronbach's alpha)

To examine internal consistency, we calculated Cronbach's alpha for the overall questionnaire and for each construct (segmental and suprasegmental features). Item-level reliability was examined using the "alpha if item deleted" function in SPSS, which showed that no item substantially decreased internal consistency. This provides preliminary evidence of reliability for the instrument.

Table 5 Cronbach's Alpha for Constructs and Items

Item-Total statistics		
	Corrected item – Total correlation	Cronbach's Alpha if item deleted
Item 1	.889	.974
Item 2	.875	.974
Item 3	.848	.975
Item 4	.864	.974
Item 5	.879	.974
Item 6	.849	.975
Item 7	.868	.974
Item 8	.896	.974
Item 9	.862	.974
Item 10	.872	.974
Item 11	.887	.974
Item 12	.857	.974

Reliability Statistics	
Cronbach's Alpha	N of Items
.976	12

Figure 4 SPSS Output Screenshot: Reliability Statistics

In sum, the questionnaire was created through a transparent and rigorous process combining theoretical grounding, expert feedback, participant-based cognitive testing, and initial psychometric checks. Collectively, these steps provide a strong justification for the questionnaire as a reliable and contextually appropriate tool for this study.

Appendix C. Interview questions

Questions about the Pros and Cons of AI Tools:

- What specific features or characteristics of AI tools do you find most beneficial for improving your pronunciation? Why?
- Are there any features or characteristics that you find frustrating or unhelpful? Could you explain?

Questions about Expectations and Suggestions for AI Tools:

- When using AI tools, what kind of feedback or features do you feel are missing but would be helpful?
- Are there any features in current AI tools that you would like to see improved? If so, how?

Questions about Attitudes toward AI tools:

1. *Do you enjoy practicing pronunciation using AI tools?*

- What specific features of AI tools make the practice enjoyable for you?
- Can you describe a moment when you felt especially engaged or uninterested while practicing?
- How would you compare this experience to practicing without AI tools?

2. *Do you feel motivated to practice your pronunciation more frequently when using AI tools?*

- What aspects of AI tools encourage you to practice more often?
- Have there been times when AI tools failed to motivate you? Why do you think that happened?
- How does your motivation change over time when using AI tools?

3. *Do you feel confident that AI tools can help you improve your pronunciation?*

- Can you recall an instance where AI tools made you feel more confident about your pronunciation?
- Are there specific feedback types from AI tools that boost or reduce your confidence? Why?
- How does this confidence compare to when you receive feedback from teachers or peers?

4. *Do you feel anxious when using AI tools to enhance your pronunciation?*

- Can you describe the kind of situations where you feel most anxious when using AI tools?
- Do you think this anxiety affects your ability to learn effectively? If yes, in what way?

5. *Do you get discouraged when AI tools point out many mistakes in your pronunciation?*

- What do you usually think or feel when the tool highlights many mistakes?
- Do AI tools often point out errors related to segmental or suprasegmental features?
- How do you react or respond to such feedback—do you keep practicing or take a break?
- Do you think this feedback style helps or hinders your progress? Why?

6. *Do you feel satisfied when you notice improvements in your pronunciation after each practice session with AI tools?*

- What specific improvements do you typically notice that make you feel satisfied?
- Do you think AI tools can improve your intonation? Why do you think so?
- Do you think AI tools can improve your sentence stress? Why do you think so?
- Do you think AI tools can improve your connected speech? Why do you think so?
- How does this satisfaction influence your willingness to continue to use these tools?