An AI-powered Chatbot to Help Non-English Majors Practice Pronunciation at a University in Hanoi, Vietnam, from the Perspectives of Students

Thuy Pham Thi Thanh[1*], Huyen Chu Thi Thu[1], Huong Nguyen Dieu[1]

[1]Hanoi National University of Education, Vietnam
[*]Corresponding author's email: phamthanhthuy09@gmail.com
[*] https://orcid.org/0000-0002-0184-1544
https://doi.org/10.54855/paic.2692

Received: 29/09/2025 Revision: 09/03/2026 Accepted: 15/05/2026 Online: 23/07/2026

ABSTRACT

Pronunciation has long been a persistent challenge for non-English majors. Although technology provides effective digital tools for practicing pronunciation, the key issue is whether learners can identify and use the most suitable tools from the outset. This study seeks to explore the design of an AI-powered chatbot to support students in practicing pronunciation, beginning with fundamental sounds. The primary research method involved consulting professional literature on AI chatbots and gathering perspectives from students and AI experts to create an effective pronunciation chatbot. The findings indicate that lecturers face considerable challenges in developing such tools due to their limited digital competence. Nevertheless, with expert collaboration, an AI pronunciation chatbot was successfully developed to help non-English majors improve their English pronunciation. The study concludes with detailed pedagogical implications for language teaching and learning.

Keywords: AI-powered chatbot, non-English majors, pronunciation, digital competence

Introduction

The mastery of English pronunciation remains a significant and long-standing obstacle for non-English majors, often hindering their ability to communicate effectively and confidently. Recent studies confirm that pronunciation difficulties among Vietnamese learners are not simply the result of a lack of tools, but stem from deeper systemic issues such as limited exposure, negative transfer from Vietnamese phonetics, and insufficient instructional support (Tran, 2024; Bui et al., 2025).

Although the digital age has introduced a wide range of language learning applications, many of these tools suffer from a "quantity over quality" paradox. They provide fragmented practice sessions that rarely translate into long-term phonological improvement. Traditional solutions such as language laboratories or basic mobile apps typically rely on standardized models and repetitive drills. While these methods can reinforce basic phoneme recognition, they often fail

CITATION | Pham, T. T. T., Chu, T. T. H., & Nguyen, D. H. (2026). An AI-powered Chatbot to Help Non-English Majors Practice Pronunciation at a University in Hanoi, Vietnam, from the Perspectives of Students. *Proceedings of the AsiaCALL International Conference*, 9, 27-39. ISSN: 2833-6836, ISBN: 979-8-90148-783-9. DOI: https://doi.org/10.54855/paic.2692

to deliver adaptive, interactive, and personalized feedback, which is essential for sustained progress (Sameera & Latha, 2024).

In response to these limitations, researchers and developers have begun exploring the potential of AI-powered chatbots. These systems embed pronunciation practice into interactive dialogue, offering real-time correction and adaptive learning pathways (Chen et el., 2024; Unite.AI, 2024). By combining the psychological benefits of conversational AI, such as reduced anxiety and increased engagement, with pedagogical depth, chatbots represent a promising step forward.

This study contributes to the ongoing debate about AI in education by shifting the focus from abstract theory to empirical application. Using a qualitative methodology that synthesizes student performance data with expert technical audits, it provides a rigorous critique of AI's classroom efficacy. The chatbot prototype developed here demonstrates how advanced AI can address the specific phonetic needs of Vietnamese non-English majors, while also helping educators overcome challenges related to limited digital competence.

As artificial intelligence becomes more deeply integrated into higher education, the central question is no longer whether AI should be adopted, but how it can be used effectively and ethically. Recent scholarship emphasizes the importance of responsible integration, highlighting both opportunities and risks (Nguyen, 2025; Wynants et al., 2025). The successful design of this chatbot not only provides a practical tool for learners but also offers broader pedagogical implications for pronunciation instruction in Vietnamese universities, underscoring the transformative potential of AI in higher education.

Literature review

This section reviews the extant scholarly literature on the core areas underpinning this study: the integration of Automatic Speech Recognition (ASR) in CALL, the efficacy of AI-powered chatbots for pronunciation training, and the specific pedagogical needs of Vietnamese learners.

Technological Integration in Pronunciation Training: From ASR to AI Chatbots

Computer-Assisted Language Learning (CALL) has increasingly adopted Automatic Speech Recognition (ASR) as a tool for pronunciation training. ASR technology provides learners with immediate feedback on their phonetic accuracy, allowing them to make quick adjustments in articulation (Fatma & Raghuvanshi, 2023; Dennis, 2022). These systems are particularly effective at identifying and correcting segmental errors, or in other words, mistakes in individual sounds. However, as Liu et al. (2025) point out, most researchers admit that ASR still struggles to evaluate suprasegmental features such as rhythm, stress, and intonation. This limitation suggests that ASR is better suited for drilling isolated sounds than for helping learners develop natural conversational fluency.

To address these shortcomings, AI-powered chatbots have emerged as a promising alternative. Unlike ASR alone, chatbots embed pronunciation practice within interactive dialogue, making the learning experience more communicative. As Chen et al. (2024) emphasized, adaptive chatbot design can personalize learning pathways, tailoring feedback to individual learners. Similarly, Kim (2019) demonstrates that chatbots can be effective in grammar instruction, showing their broader potential in language learning. Nevertheless, most chatbot systems remain "L1-agnostic," meaning they are not designed to account for the specific phonological challenges of learners from different first-language backgrounds. For example, Vietnamese learners often struggle with final consonants and stress patterns, yet these issues are rarely

addressed in globalized chatbot platforms (Tran et al., 2023). This raises questions about whether globalized chatbot platforms can adequately support learners with distinct L1 transfer issues.

Evaluating AI Chatbots for Pronunciation: Strengths and Weaknesses

AI chatbots offer several pedagogical advantages. They provide a psychologically safe environment where learners can practice without fear of embarrassment, which reduces speaking anxiety and encourages repetition. Both of them are essential for automating pronunciation skills (Kim, 2019). Their accessibility through mobile devices also aligns with the principles of Mobile-Assisted Language Learning (MALL), which promotes situated and ubiquitous learning (Kearney & Burden, 2020).

Despite these strengths, current chatbot systems often prioritize meaning over form. In practice, this means that subtle pronunciation errors may be overlooked as long as the intended word is understood. Such an approach risks fossilizing incorrect habits, particularly among L2 language beginners. Moreover, while Chen et al. (2024) highlight the potential of adaptive features, empirical evidence on their effectiveness in pronunciation training remains limited. This points to a need for more targeted research on how personalization can address specific phonological difficulties.

Evolving Pedagogies: Intelligibility and Feedback

Recent pedagogical consensus emphasizes functional intelligibility rather than native-like accuracy. Jenkins (2015), for instance, frames English as a global lingua franca, highlighting that the language is primarily used as a tool for international communication. Likewise, Derwing and Munro (2015) argue that pronunciation instruction should focus on those features most critical to comprehension, rather than attempting to eliminate every minor accent difference. This perspective permits more selective error correction, in which teachers and technological aids concentrate on high-priority issues such as stress, rhythm, or consonant clarity. By shifting the emphasis from perfecting imitation of native voice to communicative effectiveness, this approach aligns pronunciation teaching with the real demand of learners in a global context.

Another important issue in pronunciation pedagogy concerns the timing of feedback. Fu and Li (2020) demonstrate that immediate corrective feedback is more effective than delayed feedback, since it allows learners to adjust their talk in real time and reinforces accuracy and fluency more quickly. However, many ASR-driven tools still rely on automated scoring systems. This raises concerns about whether the feedback they provide truly matches pedagogical expectations, because automated scores may not capture the nuance of human communication.

Meta-analyses of MALL also show encouraging results. Tseng et al. (2022) confirm that MALL has positive effects on pronunciation development. Nevertheless, their findings also reveal considerable variation across studies. This variability suggests that outcomes are strongly influenced by contextual and learner-specific factors, such as the learners' L1 background, their level of exposure to English, and the design of the learning environment. Altogether, these insights point to the need for more nuanced approaches that adapt feedback and instructional design to the specific needs of different learners.

Context-Specific Challenges in Vietnamese EFL Acquisition

The Contrastive Analysis Hypothesis (Swan & Smith, 2001) continues to serve as a valuable framework for predicting systematic interference from the learners' first language. In the case

of Vietnamese learners, one of the most persistent difficulties involves final consonant deletion or simplification, which weakens both grammatical marking and lexical recognition (Nguyen, 2007). Furthermore, as Vietnamese is a tonal language with a syllable-timed rhythm, learners often transfer this rhythm into English, leading to errors in stress and intonation that significantly reduce intelligibility (Dang, 2017; Tran, 2023).

Despite decades of research, these challenges remain under-addressed in mainstream ASR and chatbot systems. Most local tools are designed for scalability and broad accessibility, but this emphasis often comes at the expense of localized adaptation. As a result, the phonological challenges faced by Vietnamese learners are not adequately supported. This gap highlights the need for future research to explore how adaptive AI systems can integrate insights from contrastive analysis, thereby offering feedback not only immediate but also more effective, context-sensitive, and tailored to the needs of particular learner groups.

Research Gap and Contribution

Despite the abundance of general AI tools, there is a notable scarcity of research focusing on AI chatbots tailored specifically for non-English majors at Vietnamese universities. Most existing studies focus on English majors or general adult learners, while overlooking the specific constraints faced by non-majors, such as limited curriculum time and lower baseline motivation. Furthermore, there is a lack of ASR-integrated tools that specifically target the documented Vietnamese L1-interference patterns (like final consonant deletion) in a conversational chatbot format. This gap is significant because such errors are both predictable and persistent, yet current globalized tools rarely provide tailored solutions.

The present study aims to address this shortcoming by designing a chatbot that combines the psychological benefits of AI interaction with a pedagogical focus on the specific phonetic needs of Vietnamese university students, thereby establishing a more efficient rationale for pronunciation training in Hanoi's higher education context.

Research Questions

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

(i) How can the Gemini platform be used to design and develop a pronunciation chatbot?

(ii) What is the feedback from testers regarding the effectiveness and user experience of the pronunciation chatbot?

Methods

Pedagogical Setting & Participants

The study was conducted in the real context of English teaching and learning among non-English majors at a Vietnamese university. Many students struggled to achieve standard English pronunciation, partly because they lacked solid foundational training from the beginning. Although English was taught as a compulsory school subject, their proficiency levels varied depending on their accumulated learning over time. While some managed to build relatively good language skills, others struggled to pass speaking assessments, including midterm and final exams. Pronunciation, in particular, proved to be one of the most challenging aspects for these learners, who were often unsure how to improve it effectively.

To address this issue, a pronunciation chatbot was created to support students in practicing both individual sounds and connected speech within sentences. The chatbot was developed using a

carefully designed prompt and pronunciation assessment rubrics, which provided scores as well as targeted feedback for improvement, implemented through the AI platform gemini.google.com

The prompt to create a chatbot is as below:

Role/Instruction:
You are an English pronunciation chatbot designed for Vietnamese non-English majors at VSTEP levels 3–5 (equivalent to CEFR B1–B2). You will receive the students' recordings when they read aloud a passage in English. Your task is to evaluate their pronunciation, give a score on a scale of 0–10, and provide detailed feedback.

Specific tasks:

- Listen to the student's input (assuming the system has already processed the voice into text + audio for you to analyze).
- Provide a pronunciation score:
    - Overall score from 0–10 (0 = unintelligible, 10 = accurate and natural).
    - Assess based on:
        - Segmental sounds (/θ/, /ð/, /ʃ/, …)
        - Ending sounds
        - Word stress
        - Sentence stress & intonation
        - Fluency
- Give detailed feedback in three parts:
    0. *Strengths:* highlight what the student did well.
    1. *Weaknesses:* point out specific mistakes.
    2. *Suggestions for Improvement:* suggest practical, easy-to-apply techniques (e.g., shadowing, listen-and-repeat, practicing sounds with a mirror, using online dictionaries with audio, etc.).

Feedback style:

- Concise, friendly, and encouraging.
- Use simple language, avoid overly technical terminology.
- Focus on 2–3 major errors rather than listing too many.

Final output:

- Pronunciation score (0–10).
- Feedback divided into 3 sections: Strengths – Weaknesses – Suggestions for Improvement.

Here are the rubrics for checking students' pronunciation:

Speaker's pronunciation is checked and commented on the following points (3-5 level V-step speaking assessment):

Table 1.

Marking scale for Pronunciation in Speaking skill (3-5 level V-step)

	0	1	2	3	4	5	6	7	8	9	10
Pronunciation - Individual sounds - Stress - Intonation	Test taker is not present	Performance does not satisfy band 2 descriptors	-Is often unintelligible -Can articulate a very limited repertoire of learnt words and phrases with limited accuracy	-Is mostly intelligible -Can articulate simple words and phrases but conversational partners will need to ask for repetition from time to time	-Is mostly intelligible and has acquired a quite clear pronunciation -Makes some errors with individual sounds -Shows some efforts in word stress despite frequent mispronunciations	-Is mostly intelligible and has acquired a quite clear pronunciation -Makes occasional errors with individual sounds -Shows efforts in word stress despite some mispronunciations	-Is intelligible & has acquired a quite clear & natural pronunciation -Generally clearly articulates individual sounds -Generally places word stress, but does not show efforts with sentence stress -Shows few efforts with intonation	-Is intelligible & has acquired a clear & natural pronunciation -Generally clearly articulates individual sounds -Generally places word stress, & shows efforts with sentence stress despite rather low accuracy -Shows some efforts with intonation	-Is intelligible & has acquired a very clear & natural pronunciation -Clearly articulates individual sounds -Generally places word & sentence stress -Shows good efforts with intonation	**Pro**-Is intelligible with individual sounds clearly articulated, sentence & word stress accurately placed -Has appropriate intonation -Places sentence stress flexibly & correctly to express different meanings	-Is intelligible with individual sounds clearly articulated, sentence and words stress accurately placed -Has appropriate intonation -Varies intonation and places sentence stress correctly to express different meanings and intended functions

Feedback on students' pronunciation is provided in terms of strengths, weaknesses, and areas for improvement. By reflecting on these comments, learners gain greater awareness of their pronunciation and understand how to refine it to become more accurate and expressive, approaching the naturalness of native speakers.

The participants in this study were 10 non-English majors enrolled in a General English course at a university in Hanoi, Vietnam. They were randomly selected and invited to test a pronunciation chatbot by accessing a provided link and offering feedback on their experience. The participants were between 19 and 22 years old, corresponding to first- to fourth-year undergraduate students. All participants were in good health and demonstrated active and enthusiastic engagement with the tasks assigned by the researchers. They were studying across a range of disciplines, including social sciences, natural sciences, educational sciences, information technology, arts, and foreign languages. As members of Generation Z, the students were highly familiar with digital technologies and AI tools, which they already used to support their university studies. For these reasons, the researchers placed strong confidence in the participants as suitable and reliable evaluators of the pronunciation chatbot developed for this study.

Design of the Study

Drawing on both the practical context of non-English majors' pronunciation at the university and relevant theoretical foundations, a pronunciation chatbot was developed on the Gemini platform (gemini.google.com). Theories related to pronunciation were collected and synthesized from recent, reliable sources. In addition, long-term experience in teaching pronunciation was taken into account to identify students' common strengths and weaknesses. Based on these insights, detailed guidelines for designing a pronunciation chatbot tailored to non-English majors were established. Subsequently, the chatbot was created using a specific prompt and pronunciation assessment rubrics, which provided both scores and feedback on students' actual performance. As described in the previous section, these tools formed the backbone of the system. The prompt used to develop the chatbot was revised multiple times through collaboration between researchers and AI experts specializing in artificial intelligence and AI applications in education. Following this iterative refinement process, the most reliable version of the prompt was selected to create the pronunciation chatbot employed in the present study. The chatbot was then piloted with 10 students, who provided feedback highlighting both its strengths and areas for improvement.

Data collection & analysis

The study was conducted systematically through several stages. The first step involved reviewing essential theoretical knowledge from books and journal articles on related topics. From this review, a suitable theoretical framework was adapted from previous studies to

underpin the development of a pronunciation chatbot for practice and learner experience. Drawing on these theories, an AI-driven pronunciation chatbot was built using a carefully designed prompt and pronunciation assessment rubrics (VSTEP levels 3–5), which were then implemented on the Gemini platform (gemini.google.com).

The research procedure highlighted three core tools: the prompt, the pronunciation rubric, and the Gemini AI platform. Together, they formed the foundation for creating the chatbot. The first two tools were specifically developed to reflect the practical realities of teaching and learning English pronunciation at the university level for non-English majors. They provided the structure for the chatbot, which supports students in the long-term process of mastering English pronunciation.

To analyze the outcomes, recent studies from the past few years were synthesized and compared. This was followed by an objective evaluation of the chatbot, drawing upon theoretical foundations, the prompt, and the VSTEP-based assessment rubrics. Importantly, student feedback- grounded in their actual needs for pronunciation improvement at university- was also incorporated. Such feedback offers valuable insights for refining the chatbot into a more effective version that enables students to practice pronouncing both individual words and words in connected sentences.

Findings and discussion

This is the pronunciation chatbot that was created on the Gemini platform (gemini.google.com), https://gemini.google.com/share/f53595a0d2e8

.

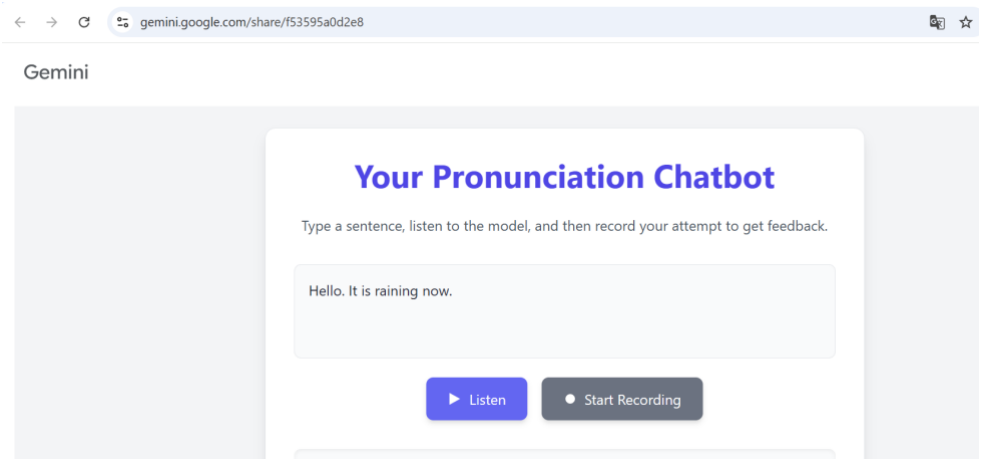

Image 1. Pronunciation chatbot

The two research questions were addressed comprehensively throughout the study. Here are the findings for the research that the authors would like to mention.

Research question (i): How can the Gemini platform be used to design and develop a pronunciation chatbot?

The pronunciation chatbot was designed and developed using a carefully constructed prompt and an assessment rubric for pronunciation (aligned with VSTEP levels 3–5), as described in the Methodology section. The detailed prompt served to guide the Gemini platform in building the chatbot. Its main function is to record a tester's spoken paragraph and provide a score, along with feedback on strengths, weaknesses, and pronunciation improvement recommendations.

Such comments were constructive, enabling testers to evaluate how well they pronounced words in sentences, recognize both strong and weak points, and identify specific areas to work on for future progress.

In addition, the pronunciation rubric (VSTEP levels 3–5) was incorporated to instruct Gemini on how to assign scores based on actual learner performance. Scores ranged from 0 to 10, reflecting the tester's pronunciation ability. Receiving a score often motivated testers, sparking curiosity about why they achieved that score and encouraging them to refine their pronunciation to attain higher scores in subsequent trials.

Together, these two tools contributed to creating a pronunciation chatbot tailored for non-English majors, aimed at strengthening their pronunciation competence to meet the requirements of the B1 VSTEP certificate- a graduation condition at university. Both tools were sufficiently reliable and specific to support the development of an effective version of the pronunciation chatbot for English learners.

Research question (ii): What is the feedback from testers regarding the effectiveness and user experience of the pronunciation chatbot?

10 non-English majors were asked to test the chatbot and give feedback after testing. Generally, good comments were collected. Here are detailed comments from them:

Concerning the chatbot's design, all participants agreed that its interface was user-friendly and well-suited for language learners. As one student remarked,

> *"The interface is neat, easy to view, and simple to use, so I quickly became familiar with it and was able to start using it without difficulty."*

The chatbot was intentionally designed to allow students to practice speaking English, record their speech, and receive immediate feedback after each attempt.

With respect to the scores and comments, all students confirmed that the system helped them clearly identify both their strengths and weaknesses in pronunciation, as well as what steps to take for improvement. One student explained,

> *"The feedback and score help me recognize which parts of my pronunciation are correct and which are not, so I can learn from mistakes, practice again, and gradually become more confident and make progress."*

Another commented,

> *"It guides me on how to improve my pronunciation and allows me to measure my progress through the scores."* An additional view highlighted novelty: *"I find it easy to learn and quite new."*

All in all, the students' feedback indicated that the scoring and evaluative comments provided meaningful guidance, enabling them to practice more effectively and steadily refine their pronunciation.

When asked about the aspects they valued most, the students expressed a range of opinions. One noted:

> *"What I like most is that the chatbot provides an immediate score after each reading, which allows me to see the outcome of every practice attempt clearly."*

Another offered a more detailed response:

> *"The feature I appreciate most is that the chatbot not only highlights the areas in my*

> *pronunciation that need improvement but also suggests ways to correct them. Moreover, it supports me with grammar and word usage. For example, when I typed and said: 'Linh is a international student in China,' the chatbot responded with: 'Article Usage and Linking: You said, "She is a international student." In English, when the next word begins with a vowel sound (like /ɪ/ in international), we use an instead of a. This is not only a grammar issue but also important for smooth, natural pronunciation, since the n sound in an connects seamlessly to the following vowel sound. Saying a international creates a slight, unnatural pause.'"* Another student remarked, *"The AI speaks with natural intonation, giving me the sense of conversing with a real person, unlike Google Translate or TFlat, which usually sound mechanical."*

Clearly, the chatbot offered unique advantages that the students recognized after several trials. These strengths should be further enhanced in the future version of the pronunciation chatbot.

Finally, regarding areas for improvement, students openly shared their suggestions. Two of them felt the current version was adequate and required no further changes. Another student commented:

> *"The chatbot's answers are somewhat general for almost any sentence, especially in terms of word stress. As for its feedback, I find it takes quite a long time to load; if the response is longer than one sentence, it becomes rather slow (possibly due to network issues). In addition, I hope the chatbot could provide feedback in Vietnamese to help learners at B1 level and below better understand their mistakes.*" A similar suggestion was: *"I hope the chatbot will provide more examples with longer sentences so that we can improve our English proficiency, and I also hope that operations (such as opening the application or scoring) will load faster."*

The authors appreciate these valuable suggestions. While loading speed largely depends on Internet connectivity and is difficult to fully resolve, the design team has already attempted to build the most specific and effective prompts possible to guide the Gemini platform. This may explain why the chatbot's comments sometimes appear repetitive across different tests. It is worth noting that testers may occasionally repeat the same pronunciation errors across multiple attempts with the chatbot, often due to individual factors. Another request from students was to receive feedback in Vietnamese. It is understandable that students were eager to fully understand the chatbot's feedback on their pronunciation, as they expected clear information about their strengths, weaknesses, and areas for improvement. This detailed feedback proved to be beneficial for their learning. Although the authors recognize this as a meaningful idea worth considering for future versions, they also believe that reading feedback in English gives students an added opportunity to practice and strengthen their English reading skills.

Discussion

Based on the findings from the two research questions above, it can be concluded that the pronunciation chatbot was carefully designed and developed step by step to support non-English majors in practicing and improving their pronunciation of both individual words and words within sentences. The Gemini platform was used to develop this tool, resulting in a relatively effective chatbot. To develop an effective prompt for the pronunciation chatbot, the researchers, together with AI experts, revised and tested multiple versions of the chatbot. The initial version allowed testers to pronounce words or sentences and receive brief feedback, but it did not include model pronunciations. The second version introduced improvements by adding performance scores and providing feedback in Vietnamese. However, the researchers

later recognized that, as students were enrolled in an English course, they could benefit more from receiving feedback in English. Consequently, a final version of the chatbot was developed that provided feedback and scores in English while retaining the target words' and sentences' model pronunciations. This refined version was considered the most effective for supporting students' independent pronunciation practice.

Overall, students' feedback was quite positive after testing it. For them, the chatbot offered a new way to practice English pronunciation. They were enthusiastic about receiving scores and comments on their spoken sentences, which enabled them to identify strengths, weaknesses, and areas for improvement. These features significantly supported their learning and are expected to help them progress toward speaking English with greater accuracy and expression. The chatbot will continue to be revised, with improved versions developed to offer students enhanced benefits and a richer language-learning experience.

Conclusion

In conclusion, the study set out to investigate the design and implementation of an AI-powered chatbot for pronunciation practice, focusing on two research questions. Regarding the first question, the findings indicate that the pronunciation chatbot was developed using *Gemini*, along with specific command-based prompts and assessment rubrics for pronunciation (VSTEP levels 3-5). These tools enabled the chatbot to provide immediate feedback on learners' pronunciation attempts, focusing initially on basic English sounds before progressing to more complex features of connected speech. Regarding the second question, tester feedback was generally positive. Students appreciated the chatbot's interactive nature, its immediate corrective responses, and the flexibility it offered for self-paced practice. Nevertheless, some testers reported technical limitations, such as slow loading of comments and a need for more advanced features to handle connected speech.

The findings of this study provide several pedagogical implications for pronunciation teaching and learning. First, language lecturers should consider integrating AI-powered chatbots into their courses as supplementary tools, particularly for non-English majors who may lack opportunities for authentic communication practice. These chatbots can provide immediate and individualized feedback, which is often difficult for teachers to deliver consistently in large classes. Second, the study highlights the importance of digital literacy among teachers. Institutions should provide professional development programs to help lecturers acquire basic technological competence so they can effectively collaborate with experts in designing and implementing AI-based tools. Training workshops on AI applications in education can empower teachers to adopt and adapt such technologies with confidence. Third, the use of chatbots should not be seen as a replacement for teacher-student interaction but rather as a complementary resource. Teachers can guide students in selecting appropriate tasks, monitor their progress, and encourage them to transfer the pronunciation skills gained through chatbot practice into real-life communication. Finally, policymakers and curriculum designers should support the inclusion of AI-enhanced pronunciation practice in language programs. By doing so, they can ensure that learners not only improve their intelligibility but also gain familiarity with modern learning technologies that are increasingly shaping the future of education.

Despite these encouraging results, the study acknowledges several limitations. It relied primarily on tester feedback and expert consultation rather than longitudinal classroom experiments. Future research should conduct larger-scale trials, incorporate diverse learner groups, and evaluate the long-term effectiveness of AI pronunciation chatbots. In conclusion,

this study provides evidence that with expert collaboration, AI-powered tools can serve as valuable support for non-English majors, helping them overcome persistent difficulties in mastering English pronunciation.

Acknowledgement

The authors would like to express their sincere gratitude to the Faculty of English, Hanoi National University of Education (HNUE) for providing the opportunity to participate in the training course *"Innovative Practices to Promote Enthusiasm and Engagement in Language Teaching"* in August 2025, in collaboration with RELO – U.S. Embassy in Hanoi.

Special thanks are extended to Dr. Tran Huong Quynh, Dean of the Faculty of English, HNUE, for her constant encouragement and support throughout the research process. We are also grateful to Dr. Greg Kessler, English Language Specialist, whose guidance and valuable insights have been instrumental in shaping the direction of this study.

In addition, we would like to thank our colleagues from the Faculties of English, Physics, Chemistry, Mathematics, and Biology at HNUE for kindly testing the chatbot and providing constructive feedback. Finally, we wish to express our sincere appreciation to the students who actively participated in the study and generously shared their experiences. Their engagement not only enriched the findings but also provided meaningful directions for further improving the chatbot.

References

Bui, T. K. P., Nguyen, T. L., Nguyen, D. Q., & Hoang, T. T. (2025). Exploring pronunciation challenges encountered by first-year non-English majored students. Hanoi University of Science and Technology, 5:67-77

https://doi.org/10.71192/jst-etu.2025.15.08

Celce-Murcia, M. (2007). Rethinking the role of pronunciation in language teaching. *Applied Linguistics, 28*(4), 507–521. https://doi.org/10.1093/applin/amm049

Chen, D.L., Aaltonen, K., Lampela, H., & Kujala, J. (2024). The design and implementation of an educational chatbot with personalized adaptive learning features for project management training. *Technology, Knowledge and Learning, 30*. https://doi.org/10.1007/s10758-024-09807-5

Dang, T. N. D. (2017). Suprasegmental errors, pronunciation instruction and communication. *Accents Asia, 9*(1), 98–120. Retrieved from https://accentsasia.org/issues/9-1/Dang.pdf

Dennis, N. K. (2022). Using AI-powered speech recognition technology to improve English pronunciation and speaking skills. *International Journal of Advanced Computer Science and Applications, 13*(9), 88–95. Retrieved from https://files.eric.ed.gov/fulltext/EJ1440171.pdf

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins. https://doi.org/10.1093/applin/amw041

Fatma, S., & Raghuvanshi, P. (2023). A review on automatic speech recognition system. *International Journal for Scientific Research and Development (IJSART), 9*(10), 91.

Retrieved from https://ijsart.com/public/storage/paper/pdf/IJSARTV9I1072441.pdf

Fu, M., & Li, S. (2020). The effects of immediate and delayed corrective feedback on L2 development. *Studies in Second Language Acquisition, 44*(1), 1–22. https://doi.org/10.1017/S0272263120000388

Jenkins, J. (2015). *Global Englishes: A resource book for students* (3rd ed.). Routledge. https://doi.org/10.4324/9781315761596

Jones, R., & Chen, L. (2021). Improving English pronunciation for non-English majors: A needs analysis. English Language Teaching Journal, 75(4), 512-526.

Liu, Y., Ab Rahman, F., & Mohamad Zain, F. (2025). A systematic literature review of research on automatic speech recognition in EFL pronunciation. *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2025.2466288

Kim, N. Y. (2019). A study on the use of artificial intelligence chatbots for improving English grammar skills. *Journal of Digital Convergence, 17*(8), 45–54. https://doi.org/10.14400/JDC.2019.17.8.45

Kearney, M., & Burden, K. (2020). Mobile learning and ubiquitous learning. In *Theorising and implementing mobile learning* (pp. 25–37). Springer. https://doi.org/10.1007/978-981-15-8277-6_3

Nguyen, K. V. (2025). The use of generative AI tools in higher education: Ethical and pedagogical principles. Journal of Academic Ethics, 23(2), 1435–1455.

https://doi.org/10.1007/s10805-025-09607-1

Nguyen, T. T. (2007). Difficulties for Vietnamese when pronouncing English: Final consonants. Retrieved from https://api.semanticscholar.org/CorpusID:43520214

Sameera, V., & Latha, M. (2024). Limitations of computer-aided language labs in schools and colleges. KLEF University. Retrieved from https://www.ijfans.org/uploads/paper/d48c01b37c48d0cb08e543be1785d22a.pdf

Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511667121

Tran, D. T. (2024). Common English pronunciation problems among Vietnamese learners: A case of non-English majors. Multidisciplinary Reviews, 7(9), e2024224. https://doi.org/10.31893/multirev.2024227

Tran, N. T., Pham, H.H., Nguyen, T.T. (2023). Exploring the challenges of L1 negative transfer among Vietnamese English language learners: A qualitative study. REFLections 31 (2), 837-857. Retrieved from https://files.eric.ed.gov/fulltext/EJ1440877.pdf

Tseng, W. T., Chen, S., Wang, S. P., Cheng, H. F., Yang, P. S., & Gao, X. A. (2022). The effects of MALL on L2 pronunciation learning: A meta-analysis. Journal of Educational Computing Research, 60(5), 1220-1252. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/07356331211058662

Wynants, S., Childers, G., De La Torre Roman, Y., Budar-Turner, D., & Vasquez, P. (2025). ETHICAL Principles AI Framework for Higher Education. California State University, Fullerton. Retrieved from https://genai.calstate.edu/communities/faculty/ethical-and-responsible-use-ai/ethical-principles-ai-framework-higher-education

Biodata

Thuy Pham Thi Thanh, Ph.D has been teaching English at Hanoi National University of Education for more than 20 years. She earned her M.A degree at the University of Languages and International Studies - Vietnam National University, Hanoi. Her Ph.D. degree focuses on linguistics in both English and Vietnamese. Her major interests are the four skills of English, vocabulary, linguistics, and methods of teaching and learning English. She has published many articles in Vietnamese and international education and linguistics journals. She has had numerous opportunities to attend various conferences, including the CamTESOL International Conference 2025 in Cambodia and the VietTESOL International Convention 2025. Her slogan is: "You cannot live the 2nd time, so live and work with love at your fullest effort!" She can be contacted at email: phamthanhthuy09@gmail.com. Orcid ID: https://orcid.org/0000-0002-0184-1544

Huyen, Chu Thi Thu, M.A, has been working as a lecturer of English at Hanoi National University of Education for 15 years. She earned her Master's degree at the University of Language and International Studies, VNU. She has a great interest in the English language as well as the cultures of English-speaking countries. She is also eager to investigate ways to improve her teaching methods and apply ICT in her teaching process. Now she is teaching subjects related to British-American cultures as well as literature at HNUE. Email: thuhuyen@hnue.edu.vn . Orcid ID: https://orcid.org/0009-0000-8451-7768

Huong Nguyen Dieu, M.A has been an English language lecturer in the Faculty of English at HNUE since 2019. Her research interests include integrating innovative skills into the EFL classroom and enhancing English language pedagogy. She is a member of the VietTESOL Masters' Group, contributing to professional development for teachers and students in Vietnam and abroad. Email: ndhuong.hnue@gmail.com . Orcid ID: https://orcid.org/0009-0006-1063-2900